



Règles de décision pour la gestion du risque : Application à la gestion hebdomadaire de la production électrique.

Romain Apparigliato

► To cite this version:

Romain Apparigliato. Règles de décision pour la gestion du risque : Application à la gestion hebdomadaire de la production électrique.. Mathématiques [math]. Ecole Polytechnique X, 2008. Français. NNT : . pastel-00004166

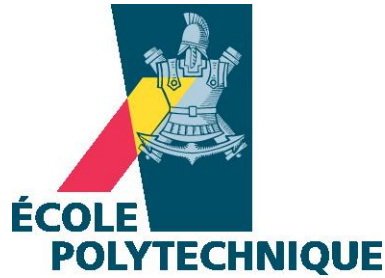
HAL Id: pastel-00004166

<https://pastel.archives-ouvertes.fr/pastel-00004166>

Submitted on 21 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Thèse présentée pour l'obtention du titre de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : Mathématiques Appliquées

par

Romain APPARIGLIATO

Titre de la thèse :

**Règles de décision pour la gestion du risque : Application à la
gestion hebdomadaire de la production électrique.**

Soutenue le 25 juin 2008 devant le jury composé de

Président :

M. NIZAR TOUZI

Rapporteurs :

M. PATRICK COMBETTES

M. PHILIPPE MAHEY

M. YVES SMEERS

Directeur de thèse :

M. FRÉDÉRIC BONNANS

Co-directeur de thèse :

M. JEAN-PHILIPPE VIAL

Invité :

M. RIADH ZORGATI

Remerciements

Mes premiers remerciements s'adressent à Jean-Philippe Vial qui a encadré cette thèse. Au delà de l'honneur de travailler avec un chercheur aussi compétent et reconnu que Jean-Philippe, j'ai eu la chance d'avoir l'encadrant tel que je l'imaginais : toujours là à me pousser quand il le fallait, d'une disponibilité, d'un enthousiasme et d'une gentillesse à toute épreuve, et favorable à ce que ces travaux soient très appliqués. Par tous ces aspects, il fut le garant de la réussite de cette thèse et je l'en remercie vivement.

Je tiens également à remercier Frédéric Bonnans qui a co-encadré cette thèse. Il fut présent lors de ces trois années et toujours d'une grande disponibilité. Par ses conseils, il a beaucoup contribué à ces travaux et à la rédaction de cette thèse. Ses nombreux encouragements sur la fin m'ont vraiment aidé à enfin achever ce manuscrit. Pour tout cela, je lui en suis très reconnaissant.

Mes remerciements vont également à Patrick Louis Combettes, Philippe Mahey et Yves Smeers pour l'intérêt qu'ils ont porté à cette thèse en acceptant d'en être les rapporteurs ; je remercie également Nizar Touzi qui a accepté de présider le jury de thèse.

Les travaux de cette thèse ont été effectués dans le cadre d'un partenariat avec EDF R&D. Je tiens donc à remercier Ryadh Boudjemadi de m'avoir fait confiance et d'avoir poussé ce projet de thèse. J'ai eu la chance d'avoir un comité de thèse qui m'a suivi très régulièrement, assurant ainsi que ces travaux soient réellement satisfaisants d'un point de vue industriel. Je remercie ainsi Xavier Warin qui pilote le projet responsable de cette thèse. Ses connaissances notamment en statistiques et en finance furent d'une aide très précieuse ! Il m'a également laissé participer à de nombreuses conférences ce qui m'a permis assurément de m'enrichir et de compléter ma formation auprès des plus grands chercheurs en optimisation stochastique. Je tiens ensuite à remercier Gabriele Ripper qui m'a accompagné durant ma première année. Elle a permis de mettre cette thèse sur les rails notamment par sa grande rigueur. Son enthousiasme, sa gentillesse et ses chocolats nous ont beaucoup manqué... Enfin, au même titre que Jean-Philippe, je tiens à remercier le plus chaleureusement possible Riadh Zorgati. Ce fut un grand plaisir de travailler ensemble. Toujours disponible et à l'écoute, des idées à revendre et tous ses commentaires, en particulier sur le manuscrit, furent très précieux. Que de bons souvenirs à refaire le monde durant ces heures de discussion, à rigoler aux jeux de mots les plus tordus jamais entendus jusque là...

J'ai également une pensée pour tous les membres du département OSIRIS dont j'ai eu le plaisir de côtoyer durant ces trois années. Les pauses café et les discussions de bout de couloir furent des moments de détente appréciables. Je tiens à féliciter mes nombreux collègues de bureau pour avoir réussi à me supporter : Grégory (deux fois !), Rémi, Benjamin (une semaine...), Agnès, Nazim et la record-woman Sylvie (plus d'un an !). Je remercie le groupe R37 de m'avoir accueilli à mon arrivée et de penser à moi pour les repas quand j'étais seul dans mon couloir, et R32 pour cette dernière année pleine d'ambiance et de soirées Bureau ou bowling. Merci à mes voisines de bureau qui ont su rendre joyeuses les pauses M&M's (Chloé, Sylvie, Lise, Steph,...) ! Je remercie

aussi tous les footballeurs qui ont participé à l'aventure des Pharaons. On n'était pas vraiment bons mais au moins on rigolait bien... Merci à mes camarades du pôle d'optimisation stochastique pour les bonnes rigolades en conférence ou aux soirées poker : Kengy, Arnaud, Pierre, Cyrille, Grégory,... Enfin, je remercie toutes les autres personnes du département -que je ne nommerais pas pour n'oublier personne- qui ont fait que ces trois années furent une superbe expérience.

Pendant ces trois ans, j'ai aussi eu l'occasion de séjourner régulièrement au sein du Logilab à HEC Genève. Je tiens à remercier tous ses membres pour leur accueil, en particulier Alain, Julien, Frédéric et Laurent. Ce fut toujours avec grand plaisir que je les retrouvais. Je tiens d'ailleurs à dire un gros merci à Julien Thénier avec qui j'ai beaucoup travaillé, en particulier sur l'application de la SPSDR à mes problématiques de gestion de risque en électricité. J'ai pu bénéficier de toute son expérience sur la méthode et ainsi bénéficié d'un gain de temps certain. J'aurais certainement eu beaucoup de difficultés à arriver au bout des expérimentations numériques sans cette aide... Et puis ses conseils en informatique sont toujours précieux !

J'ai aussi envie de partager ce moment avec mes amis custards, toujours là quand il faut. Merci à vous Eric, Sylvain, Mathieu. Sans oublier évidemment Isabelle et la petite Pauline (en attendant la petite dernière)...

Je suis aussi très reconnaissant envers Delphine Sinoquet, sans qui je n'aurais peut-être pas fait cette thèse. Elle a su, au cours de mon stage de DEA, me donner pleinement confiance en mes moyens et me donner envie de faire de la Recherche en R&D.

Après EDF, j'ai finalement choisi de rejoindre l'équipe du Pôle Simulation et Optimisation de Gaz de France. Je remercie Fabrice Chauvet de m'avoir fait confiance et de m'avoir permis de terminer ma thèse dans les meilleures conditions possibles. Je remercie aussi tous les membres de l'équipe pour leurs encouragements et pour leurs nombreux conseils lors de mes répétitions. Enfin un grand merci au Maître David qui m'a accordé beaucoup de son temps. Même si à cause de lui les dernières nuits furent courtes, cette aide fut très précieuse...

Un grand merci à mes parents, pour leurs encouragements et leur soutien depuis mon enfance. Il est évident que sans eux, je n'aurais jamais réussi à arriver jusque là. J'ai d'ailleurs une pensée émue pour mon père qui aurait souhaité vivre ce moment... Enfin, c'est vers Aurélie que je me tourne. Je lui dédie ce mémoire pour m'avoir supporté pendant ces trois années, m'avoir encouragé, avoir enduré mes horaires tardifs et mes sautes d'humeur : bref, merci pour tout !

Organisation de la thèse

Le thème général de cette thèse est la gestion du risque de défaillance physique dans le cadre de la planification de la production d'électricité à l'horizon hebdomadaire. Il s'agit d'un problème difficile en raison de la taille importante du parc français de production, de sa grande diversité (et donc de contraintes techniques) et du contexte d'incertitude assez fort (plusieurs aléas à forte variabilité comme la demande, les prix de marché, l'hydraulicité, les pannes de centrales,...). Jusqu'à présent, ce dernier point n'est pas intégré aux outils de cet horizon. Nous concentrerons donc nos efforts sur l'étude et l'application de méthodes d'optimisation dans l'incertain.

Tous ces travaux ont été effectués en collaboration avec Jean-Philippe Vial (HEC Genève, OR-DECSYS), Riadh Zorgati (Département Optimisation, SIMulation, RISque et Statistiques appliqués aux marchés de l'énergie, EDF R&D), et J.-Frédéric Bonnans (Centre de Mathématiques Appliquées de l'Ecole Polytechnique).

La première partie de la thèse est découpée en deux chapitres d'introduction générale.

- Le chapitre 1 introduit la problématique de la gestion des risques physiques chez un producteur d'électricité. Pour cela, nous présentons la procédure de gestion de la production d'électricité à EDF et en particulier le découpage temporel des décisions et des flux d'informations à transmettre d'une période à l'autre. Nous insistons sur le rôle d'un horizon étudié récemment : l'hebdomadaire (12 jours). Cet horizon a un rôle particulier puisqu'il est : d'une part, la "sentinelle" de l'horizon journalier en simulant l'état du parc au vu des informations fournies par les horizons antérieurs et, d'autre part, l'horizon sur lequel des décisions de sécurité peuvent être prises afin de s'assurer que le risque de défaillance physique soit réduit. Pour maximiser les gains économiques associés à chacun de ces rôles, il est très important de savoir prendre en compte les aléas dans les outils dédiés à cet horizon. Après cette introduction générale au contexte, une synthèse détaillée des travaux de la thèse est présentée.
- Le chapitre 2 est un rapide survol de l'état de l'art en optimisation dans l'incertain. Nous présentons les principales approches existantes afin de situer le cadre théorique dans lequel cette thèse s'inscrit. Nous différencions les problèmes à 2 étapes, où l'aléa de tout l'horizon est dévoilé dès le premier pas de temps, des problèmes plus complexes à plusieurs étapes où l'aléa n'est dévoilé qu'à chaque instant. Ce type de problème est très difficile et il n'existe aucune approche convaincante dans la littérature pour les résoudre efficacement. Deux pistes semblent cependant intéressantes et seront explorées dans la suite : l'optimisation robuste et la programmation stochastique avec règles de décision constantes par morceaux.

La deuxième partie est consacrée à l'étude de l'optimisation robuste et à son application sur un problème concernant le premier rôle de l'horizon hebdomadaire. Elle est également constituée de deux chapitres.

- Le chapitre 3 est une introduction à l'optimisation robuste linéaire. Nous effectuons un état de l'art de cette méthode, restreint au cadre linéaire pour que la présentation soit la plus claire possible. Nous montrons notamment comment faire le lien entre l'optimisation robuste et les

contraintes en probabilité (généralement non convexes voire non connexes). Nous décrivons ensuite comment obtenir l'équivalent robuste d'une contrainte d'inégalité dans le cas d'ensembles d'incertitude issues de normes simples (ℓ_1 , ℓ_2 , ℓ_∞). Nous généralisons cette construction au cas d'un ensemble d'incertitude polyédrique. Pour mesurer la qualité de l'approximation d'une contrainte en probabilité par l'optimisation robuste, nous présentons et améliorons quelques résultats permettant de borner la probabilité de satisfaction de la contrainte pour une solution robuste dans le cas des normes usuelles. Nous proposons ensuite des extensions à une telle approche. Nous citons notamment : l'extension au cas multi-périodique via des règles de décision linéaires en l'aléa, et la robustesse étendue permettant de contrôler les violations (en pénalisant leur distance à l'ensemble d'incertitude).

- Le chapitre 4 applique l'optimisation robuste au problème de la gestion hebdomadaire d'une vallée hydraulique. Cette problématique est très intéressante car l'hydraulique, étant disponible très rapidement et ne coûtant rien, est une ressource de substitution aux unités chères lors de pointes de consommation. Une gestion efficace peut donc permettre de dégager des gains assez importants. La forte variabilité des apports naturels (pluie, fonte des neiges,...) la rend toutefois difficile. Les réservoirs devant satisfaire des contraintes légales de niveaux de volume, les apports peuvent contraindre les exploitants à corriger les plannings en urgence à vue d'expert (sans garantie d'optimalité). Une gestion robuste par rapport à cet aléa permet d'obtenir des commandes moins sujettes à être corrigées par des exploitants et donc plus performantes économiquement. Nous avons testé diverses approches robustes (une version statique avec des révisions périodiques, une version avec règles de décision). Nous comparons les résultats avec un modèle Déterministe à Révisions Périodiques (DRP), devant simuler le comportement d'un exploitant de vallée. Nous montrons que nous pouvons quasiment annuler les violations de contraintes par rapport au DRP avec une augmentation de coût de l'ordre de seulement 0,6% par rapport à la borne minimale qu'est le coût en information parfaite.

La troisième partie de cette thèse est consacrée au deuxième rôle de l'horizon hebdomadaire : la gestion du risque de défaillance physique. Elle est découpée en trois chapitres.

- Le chapitre 5 présente la problématique générale. Il s'agit de prendre des décisions de sécurité (modification du parc de production, achats sur les marchés à terme, déclenchement de contrats d'effacement avec des clients) pour que le risque 1% soit satisfait, c'est à dire pour qu'il n'y ait qu'un scénario sur 100 possibles sur lequel la demande ne soit pas satisfaite. Ce critère est en fait une formulation simple des dispositions réglementaires auxquelles EDF doit faire face. Actuellement, ce processus est effectué après avoir optimisé le parc de production (en déterministe) et sur la connaissance de la marge de production (écart entre l'offre totale et la demande). A partir d'un calcul de la marge (les centrales sont toutes disponibles à leur maximum à tout instant et les aléas sont gaussiens), les gestionnaires prennent des décisions réalisables. Dans ce chapitre, nous étudions la possibilité d'intégrer le risque 1% directement dans le problème de gestion de production. Ce problème étant trop difficile, nous le simplifions pour finalement justifier l'approche appliquée actuellement par les gestionnaires. Nous formulons toutefois le problème sous la forme d'un problème linéaire en nombres entiers permettant de calculer des solutions plus performantes économiquement. Pour cette optimisation, nous supposons avoir à disposition des scénarios de marge et de prix de marché, et nous réduisons les décisions aux achats (et ventes) sur les marchés et aux contrats d'effacement. Enfin, nous

nous plaçons dans un cadre statique où les aléas sont dévoilés sur toute la période dès l’instant initial.

- Le chapitre 6 est un chapitre d’exploration consacré à l’extension du problème précédent au cas multipériodique (l’aléa est dévoilé à chaque instant). Ce problème est très difficile, notamment à cause du risque 1% formulé comme une contrainte en probabilité. Pour le résoudre, nous appliquons la programmation stochastique avec règles de décision constantes par morceaux. Il s’agit d’une heuristique qui, en relâchant la contrainte de raffinement de la programmation stochastique sur arbre classique, permet de contrôler la taille de l’arbre. Nous présentons donc dans ce chapitre cette approche originale, ainsi que les extensions (cas d’un aléa multidimensionnel, contraintes sur les états) mises en place dans cette thèse. Nous l’appliquons à une version simplifiée du problème de couverture précédent. Les résultats obtenus sont prometteurs et montrent que, pour certains jeux de paramètres, nous pouvons obtenir des règles permettant de dégager des bénéfices supérieurs à ceux obtenus par des approches statiques (modèle avec contraintes en probabilité et robuste).
- Le chapitre 7 décrit la génération des scénarios de marge et de prix (Spot et Week Ahead) nécessaires pour les modèles des deux chapitres précédents. La marge est calculée actuellement en faisant un calcul simpliste basé sur l’hypothèse d’aléas gaussiens. Nous proposons une approche plus réaliste pour améliorer ce calcul. La capacité maximale hydraulique est calculée pour chaque vallée, à l’aide d’un signal de prix construit de sorte à ce que la vallée produise à son maximum à chaque heure de pointe. La capacité maximale obtenue est ainsi réaliste car vérifiant les contraintes de gestion de la vallée. Ce calcul est effectué pour différents scénarios d’apports hydrauliques. La marge thermique est calculée en prenant la somme des productions maximales des centrales disponibles en chaque instant. L’aléa de panne est introduit à l’aide d’une loi exponentielle. La demande est générée à l’aide d’un modèle auto-régressif complexe, fourni par le groupe de “Prévision de consommation” d’EDF. Les prix de marché sont générés à l’aide d’un modèle classique à 2 facteurs. Il a été fourni par le groupe “Risque et Valorisation” d’EDF. Nous décrivons brièvement ce modèle.

Enfin, les chapitres de conclusions et de perspectives constituent la dernière partie de cette thèse.

- Le chapitre 8 propose un rapide résumé de la thèse basé sur les trois axes de cette thèse : appropriation et application d’approches originales d’optimisation dans l’incertain, modélisation de problèmes (en particulier de risque) et simulation des solutions pour valider une approche.
- Le chapitre 9 propose enfin un large éventail de perspectives recouvrant l’ensemble des problématiques étudiées dans cette thèse. Elles sont soit fondées sur des réflexions, soit basées sur des modèles initiés mais non aboutis faute de temps.

Nous avons essayé de faire en sorte que chaque chapitre de la thèse puisse être lu indépendamment des autres. En particulier, une bibliographie clôture chacun d’entre eux.

Table des matières

I	Introduction	15
1	Introduction	17
1.1	Contexte	18
1.1.1	Introduction générale	18
1.1.2	Le Parc de Production	19
1.1.3	La gestion de la production électrique	20
1.1.4	Emergence d'une gestion hebdomadaire/mensuelle de la production . . .	22
1.1.5	Sujet de thèse	23
1.2	Contexte : méthodes d'optimisation utilisées	23
1.2.1	L'optimisation robuste	23
1.2.2	La programmation stochastique avec règles de décision constantes par morceaux	25
1.3	Synthèse des résultats	28
1.3.1	La gestion robuste d'une vallée hydraulique	28
1.3.2	La gestion du risque physique	33
1.4	Pour résumer : Apports principaux de la thèse	41
1.4.1	Aspects méthodologiques de la thèse	41
1.4.2	Modélisation de problèmes industriels	42
1.4.3	La validation par simulation	44
1.5	Publications et communications orales	45
1.5.1	Communications orales	45
1.5.2	Articles et chapitres de livre	45
1.5.3	Notes internes	46
	Bibliographie	46
2	L'optimisation dans l'incertain	51
2.1	Les modèles statiques	52
2.1.1	Le gradient stochastique	53
2.1.2	L'approche par scénarios	54
2.1.3	L'aléa dans les contraintes	54
2.2	Les problèmes à deux étapes	56
2.3	Les problèmes à plusieurs étapes	60
2.3.1	La programmation stochastique sur arbre	60
2.3.2	La programmation dynamique et ses extensions	62

TABLE DES MATIÈRES

2.3.3	Les autres approches	64
2.4	Conclusions	64
	Bibliographie	65
II	Etude et application de l'optimisation robuste	73
3	Introduction à l'optimisation robuste linéaire	75
3.1	Introduction	76
3.2	Problème d'optimisation linéaire avec coefficients incertains	78
3.3	Satisfaction en probabilité d'une contrainte d'inégalité	79
3.3.1	Distribution normale	80
3.3.2	Distribution uniforme	80
3.3.3	Distribution générale	82
3.4	Solution robuste pour un ensemble d'incertitude	83
3.4.1	Contrainte robuste	84
3.4.2	Equivalent robuste d'une contrainte d'inégalité	86
3.5	Probabilité de satisfaction de la contrainte pour une solution robuste	90
3.5.1	Ensemble d'incertitude $\Xi = B(0, k)_2$	91
3.5.2	Ensemble d'incertitude $\Xi = B(0, 1)_\infty \cap B(0, k)_1$	93
3.6	Eléments pour définir l'ensemble d'incertitude	94
3.6.1	Une approche heuristique basée sur un aléa gaussien	94
3.6.2	Equivalent robuste	95
3.7	Problème général d'optimisation linéaire	97
3.7.1	Equivalent robuste d'une contrainte d'égalité	97
3.7.2	Performance moyenne et performance garantie	98
3.8	Problèmes dynamiques	99
3.8.1	Problèmes avec recours fixe	101
3.8.2	Règle de décision linéaire par morceaux avec coefficients de recours fixes	101
3.8.3	Problèmes avec coefficient de recours incertain	102
3.9	Extension du concept de robustesse : violation contrôlée des contraintes	103
3.9.1	Définition formelle de la robustesse étendue	103
3.9.2	Robustesse étendue dans le cadre de la programmation linéaire	104
3.10	Conclusion	107
	Bibliographie	108
4	L'optimisation robuste appliquée à la gestion court-terme d'une vallée hydraulique	111
4.1	Introduction	112
4.2	Modèle déterministe de gestion d'une vallée hydraulique	114
4.2.1	Modélisation choisie	114
4.2.2	Modèle déterministe	116
4.3	Incertitude	117
4.3.1	Incertitude sur les apports d'eau	117
4.3.2	Les recours	119
4.3.3	Optimisation, incertitude et validation	119

TABLE DES MATIÈRES

4.4	Politique de gestion Robuste à Révisions Périodiques (RRP) du problème hydraulique	121
4.5	Politique de gestion Robuste à Ajustements Linéaires (RAL) du problème hydraulique	122
4.5.1	Contraintes de turbinage	124
4.5.2	Contraintes de volume	125
4.5.3	Fonction-objectif	127
4.6	Tailles des modèles	127
4.7	Etude expérimentale	130
4.7.1	Implémentation	130
4.7.2	Jeu d'essais	131
4.7.3	Simulation	132
4.8	Expérimentations numériques	135
4.8.1	Comparaison entre DF et DRP	135
4.8.2	Comparaison entre DRP et RRP	136
4.8.3	Comparaison entre RRP et RAL	140
4.9	Expérimentations numériques : débit minimum	144
4.10	Conclusion	145
	Bibliographie	146
III	Gestion de la couverture contre le risque de défaillance physique	149
5	La gestion du risque physique	151
5.1	Introduction à la gestion du risque physique	152
5.1.1	Obligation réglementaire de gestion du risque de défaillance physique . . .	152
5.1.2	Procédure actuelle de couverture, limitations	152
5.1.3	Les outils financiers disponibles sur les marchés	153
5.2	Etude de l'existant à EDF	154
5.3	Une approche opérationnelle simplifiée	159
5.4	Modélisation du problème en boucle ouverte	163
5.4.1	Boucle ouverte et boucle fermée	163
5.4.2	Formulation du problème avec les actifs financiers	164
5.4.3	Calcul des quantiles	170
5.4.4	Résultats	170
5.5	Conclusion	171
	Bibliographie	171
6	Boucle fermée : la programmation stochastique avec Règles de Décision constantes par morceaux	173
6.1	Introduction	174
6.2	Programmation stochastique avec règles de décision constantes par morceaux . .	176
6.2.1	Formalisme	176
6.2.2	De la programmation stochastique à la SPSDR...	177
6.2.3	L'approche SPSDR en détails	181
6.2.4	Résultats obtenus sur le problème de chaînes d'approvisionnement	186

TABLE DES MATIÈRES

6.3	Adaptations de l'approche pour le problème de couverture	186
6.3.1	Aléa multidimensionnel	186
6.3.2	Contraintes sur l'état	188
6.3.3	Comparaisons numériques entre les heuristiques de calcul du problème p -médianes	188
6.4	Formulation simplifiée du problème	190
6.5	Ecriture du problème de couverture avec SPSDR	192
6.6	Résultats numériques : pré-requis	193
6.6.1	Jeu d'essais	193
6.6.2	Modèles simples de comparaison	194
6.7	Résultats numériques avec SPSDR	197
6.7.1	Effets du nombre d'ensembles de décisions et de la taille des experts . . .	198
6.7.2	Effets du nombre d'experts	200
6.7.3	Introduction d'un coefficient d'immunité	201
6.7.4	Synthèse des résultats	203
	Bibliographie	205
7	Les processus stochastiques du problème de couverture	207
7.1	Introduction	208
7.2	La marge physique de production	209
7.2.1	Définition	209
7.2.2	Estimation de la marge avec des aléas gaussiens	210
7.2.3	Estimation de l'offre de production thermique	211
7.2.4	Estimation de l'offre de production hydraulique	214
7.2.5	Estimation probabiliste de la marge	220
7.2.6	Résultat typique	222
7.2.7	Estimation de modèles ARMA	223
7.3	Les prix de marché	224
7.3.1	Introduction	224
7.3.2	Définitions	225
7.3.3	Du modèle factoriel unitaire au modèle factoriel par produit	226
7.3.4	Discrétisation en temps du modèle à 2 facteurs "par produit"	230
7.3.5	Approximation sur les facteurs de forme	231
7.3.6	Création de scénarios	232
	Bibliographie	232
IV	Conclusions et perspectives	235
8	Conclusions	237
9	Quelques perspectives pour l'après thèse	239
9.1	Perspectives sur la gestion robuste d'une vallée hydraulique	240
9.1.1	Pistes sur l'application de l'optimisation robuste	240
9.1.2	Pistes sur la modélisation du problème	241
9.2	Perspectives sur la gestion des risques de défaillance physique	241

9.2.1	Le calcul de la marge	242
9.2.2	Boucle ouverte : calcul de marge intégré au problème de couverture . . .	243
9.2.3	Boucle fermée	250
V	Annexes	257
A	Application de l'optimisation robuste à la gestion hydraulique hebdomadaire : compléments	259
A.1	Equivalent robuste de l'objectif	259
A.2	Etude du comportement du modèle robuste (type RAL) selon la couverture de l'objectif	262
A.3	Etude du comportement du modèle robuste (type RAL) selon le nombre de termes dans la règle de décision	264
A.4	Comportement du modèle RAL selon le niveau de robustesse des contraintes . .	267
B	Comparatif entre les distributions de marge empiriques et une approximation par loi normale	271
C	Calcul de la marge physique de production : calculs complémentaires	273
C.1	Capacité de production maximale thermique : calcul des probabilités de panne .	273
C.1.1	Phase 1 : initialisation	274
C.1.2	Phase 2 : généralisation	275
C.2	Capacité de production maximale hydraulique : calcul des paramètres de la loi normale bi-tronquée	278
D	Estimation de la densité par la méthode à noyau	281
E	Estimation de la marge à partir de scénarios	283
E.1	Distributions de marge	283
E.2	Comparaison des valeurs de quantiles pour 5 000, 7 500 et 75 000 scénarios . . .	287

TABLE DES MATIÈRES

Première partie

Introduction

Chapitre 1

Introduction

Sommaire

1.1	Contexte	18
1.1.1	Introduction générale	18
1.1.2	Le Parc de Production	19
1.1.3	La gestion de la production électrique	20
1.1.4	Emergence d'une gestion hebdomadaire/mensuelle de la production	22
1.1.5	Sujet de thèse	23
1.2	Contexte : méthodes d'optimisation utilisées	23
1.2.1	L'optimisation robuste	23
1.2.2	La programmation stochastique avec règles de décision constantes par morceaux	25
1.3	Synthèse des résultats	28
1.3.1	La gestion robuste d'une vallée hydraulique	28
1.3.2	La gestion du risque physique	33
1.4	Pour résumer : Apports principaux de la thèse	41
1.4.1	Aspects méthodologiques de la thèse	41
1.4.2	Modélisation de problèmes industriels	42
1.4.3	La validation par simulation	44
1.5	Publications et communications orales	45
1.5.1	Communications orales	45
1.5.2	Articles et chapitres de livre	45
1.5.3	Notes internes	46
	Bibliographie	46

Ce chapitre introduit le contexte général de la thèse et l'ensemble des travaux effectués. Dans un premier temps, nous décrivons la problématique industrielle ainsi que les méthodes d'optimisation dans l'incertain appliquées. Puis, dans un deuxième temps, nous effectuons une synthèse détaillée des travaux accomplis tout au long de cette thèse.

1.1 Contexte

Dans cette section, nous présentons le contexte général dans lequel s'inscrit cette thèse. Après une introduction très générale, nous introduisons les problématiques de la gestion de production et de la gestion du risque physique de défaillance à EDF.

1.1.1 Introduction générale

A l'origine, dans l'Union Européenne, l'industrie électrique (tout comme les autres industries de réseau comme le gaz) était réglementée et intégrée sur la chaîne de valeur. Les coûts d'investissement nécessaires en terme d'infrastructures sont en effet si élevés que les entreprises ne pouvaient se risquer à les prendre en charge sans une certaine garantie des autorités. Cependant, le développement des théories industrielles des années 80 a eu tendance à montrer que ces monopoles intégrés n'étaient pas incités, de par leur position, à être les plus compétitifs possibles. Ces théories suggèrent de réintroduire la concurrence aux niveaux des activités où cela est applicable (production et négoce). A l'inverse, les activités de transport et de distribution, activités de monopole naturel caractérisées par des rendements d'échelle (les coûts moyens diminuent quand les quantités fournies augmentent), restent sous contrôle actif du régulateur.

Avec cette libéralisation du secteur énergétique et l'ouverture des marchés de l'énergie, la gestion économique de chaque producteur a ainsi évolué. Auparavant, la fourniture d'énergie à chaque client et à tout moment était prioritaire et s'effectuait à moindre coût. Mais l'ouverture à la concurrence oblige chaque producteur à être désormais compétitif, tout en continuant évidemment de satisfaire ses obligations de fourniture. Dans ce contexte, l'apparition des marchés joue un double rôle. Ils permettent, à l'aide d'achats (ou de ventes) de sécurité, d'optimiser la couverture de l'entreprise contre le risque de non-satisfaction de la demande de ses clients. Mais, ils peuvent aussi permettre de tirer avantage de situations particulières (prix de marché avantageux) et ainsi dégager des bénéfices financiers supplémentaires.

Cependant, la prise en compte de l'achat sur les marchés de contrats de fourniture d'électricité comme nouvel outil de couverture conduit aussi à revoir la problématique de la couverture physique. Celle-ci consiste à déterminer quelles décisions optimales, selon un certain critère économique, permettent de se couvrir contre un risque trop élevé de défaillance physique. Il s'agit des situations où le producteur n'est pas capable, pour diverses raisons (conditions climatiques extrêmes entraînant une demande accrue, indisponibilité d'unités de production, mauvaise anticipation de certains phénomènes, ...), de satisfaire sa demande en énergie. Pour évaluer ce risque, chaque producteur définit son propre critère tout en contrôlant le coût de ces actions de couverture (pour éviter des surcoûts excessifs, en particulier sur les marchés). Cette problématique de couverture globale lie donc deux types de couverture : la couverture physique, utilisée pour

réduire le risque physique de défaillance et la couverture financière, mise en œuvre pour réduire le risque de pertes financières excessives.

Ces risques apparaissent principalement car la gestion d'un parc de production électrique est soumis à de nombreux aléas rendant ainsi la planification difficile à élaborer. Il peut s'agir des aléas de consommation, d'indisponibilité des différentes unités de production, d'hydraulicité, de prix de marché,... Pour réduire ces risques, il est donc nécessaire de savoir prendre en compte au mieux ces aléas au sein même de la planification. Nous devons donc faire face à un problème d'optimisation dans l'incertain. Il existe différentes approches basées sur des hypothèses différentes : connaissance probabiliste des aléas ou non, optimisation sur un critère économique en espérance ou non,... Les approches classiquement appliquées pour ce type de problèmes souffrent de différents inconvénients : nécessité d'une distribution probabiliste des aléas et taille des problèmes dans le cadre de la programmation stochastique, nombre de stocks ou de pas de temps dans les approches de type programmation dynamique, calcul d'espérances mathématiques multidimensionnelles,... Cet état de fait nous conduit à étudier de nouvelles méthodologies.

1.1.2 Le Parc de Production

Pour produire de l'électricité, EDF dispose d'un parc d'actifs physiques et financiers conséquent et hétérogène.

Les actifs physiques

Le parc thermique est composé de 58 tranches nucléaires et de 50 unités thermiques classiques (charbon, houille, fioul et turbines à gaz). Pour chaque groupe thermique, sont donnés :

- un ensemble de contraintes à respecter, tant statiques que dynamiques telles que la plage et les paliers de production, les impositions de puissance, les durées minimales de fonctionnement, d'arrêt, les paliers de démarrage, etc.
- une structure de coût : coût de démarrage, coût fixe, coût proportionnel ; Ainsi, le coût de démarrage est une fonction de l'arrêt qui précède (refroidissement d'une centrale) alors que le coût de fonctionnement est une fonction affine de la puissance produite (coûts fixes et différentiels ou proportionnels).

Le parc hydraulique est organisé en 20 vallées nationales, chaque vallée étant modélisée comme un ensemble cohérent de réserves hydrauliques et d'usines de production, interconnectées. Il représente plus de 100 usines hydrauliques. Chaque vallée est caractérisée par sa typologie (cf figure 1.3.1) qui décrit les interconnexions entre réserves et usines, leur nature (usine de pompage ou de turbinage) et certains paramètres comme le délai de parcours de l'eau entre deux réservoirs. Pour chaque usine et pour chaque réserve hydrauliques, sont donnés :

- un ensemble de contraintes à respecter telles que les volumes d'eau minimaux et maximaux des réservoirs, les limitations de turbinés et/ou de puissance, etc.
- une structure de coût définie à partir d'une valeur d'usage ω de l'eau, calculée par des techniques de gestion de stock à l'horizon annuel ; la valeur d'usage de l'eau représentant l'économie

réalisée sur un planning de production thermique en utilisant l'eau au meilleur moment. Il s'agit donc d'une valeur de substitution.

- un ensemble de courbes débit/puissance permettant de déterminer la puissance obtenue pour un débit donné, compte tenu du rendement de l'usine et de son niveau de production (courbe linéaire par morceaux).

Les actifs financiers

EDF utilise la tarification comme levier de modulation de la courbe de charge afin de mieux répartir la consommation dans le temps. Ceci se traduit par la définition de contrats passés avec des clients. Ces contrats tarifaires (comme les Effacements Jours de Pointe ou EJP) sont des contrats présentant deux types de tarification : une tarification plus faible que la moyenne la grande majorité du temps et une tarification beaucoup plus élevée sur un nombre de jours donné. Ces jours sont choisis un à un par EDF entre novembre et mars inclus, avec un préavis de quelques heures. Lors d'une de ces journées, les clients concernés préfèrent logiquement diminuer leur consommation électrique plutôt que payer un prix très élevé. Cette quantité d'énergie effaçable peut être vue comme un stock supplémentaire d'énergie disponible. Globalement, nous estimons à 5000MW la quantité d'énergie disponible.

1.1.3 La gestion de la production électrique

Pour utiliser au mieux les grands stocks (Hydraulique et Nucléaire) et pour assurer un coût de production global minimal, il est nécessaire d'optimiser le planning de production. Cependant, ce problème d'optimisation doit prendre en compte des décisions concernant divers horizons de temps : des investissements au pluriannuel jusqu'aux ajustements de production à l'horizon infra-journalier. Résoudre ce problème de façon globale est évidemment trop difficile et a entraîné un découpage du problème en autant de problèmes que d'horizons de temps considérés. Cette approche est sous-optimale mais est, jusqu'à présent, nécessaire. Les décisions pour chaque horizon sont les suivantes :

- **Long terme** (10 à 20 ans) : investissements (renouvellement du parc, nouvelles technologies, types de centrales...);
- **Pluriannuel** (5 ans) : planification des arrêts des tranches nucléaires pour rechargement du combustible nucléaire;
- **Annuel** (1 an) : calcul des valeurs d'usage des grands stocks hydraulique et nucléaire, évaluation du risque de défaillance physique, définition de la stratégie de passage de l'hiver prochain, négociation si nécessaire des contrats d'approvisionnement complémentaires auprès d'autres producteurs ou de traders;
- **Hebdomadaire** (7 à 12 jours) : simulation du parc de production pour le journalier, placements des essais pour maintenance des centrales, couverture à court-terme et propositions de déclenchements d'options tarifaires;
- **Journalier** (1 à 2 jours) : détermination du programme de marche pour RTE¹, déclenchements

¹Réseau de Transport Electrique. Il s'agit du gestionnaire du réseau devant : gérer les infrastructures du réseau, gérer les flux d'électricité sur les réseaux et contribuer au bon fonctionnement du marché de l'électricité. Il assure

1.1. CONTEXTE

d'options tarifaires ;

- **Infrajournalier** (quelques heures) : adapter quasiment en continu les plannings de production à l'évolution réelle de la demande, proposer à RTE des offres d'ajustement.

EDF a donc dû développer des méthodes et des outils de gestion de la production à chacun de ces différents horizons en vue de rendre traitable le problème global de la gestion de la production, et donc réduire les coûts de production. Ces divers horizons de temps sont caractérisés par des décisions propres et par des indicateurs particuliers liant les horizons entre eux (figure. 1.1).

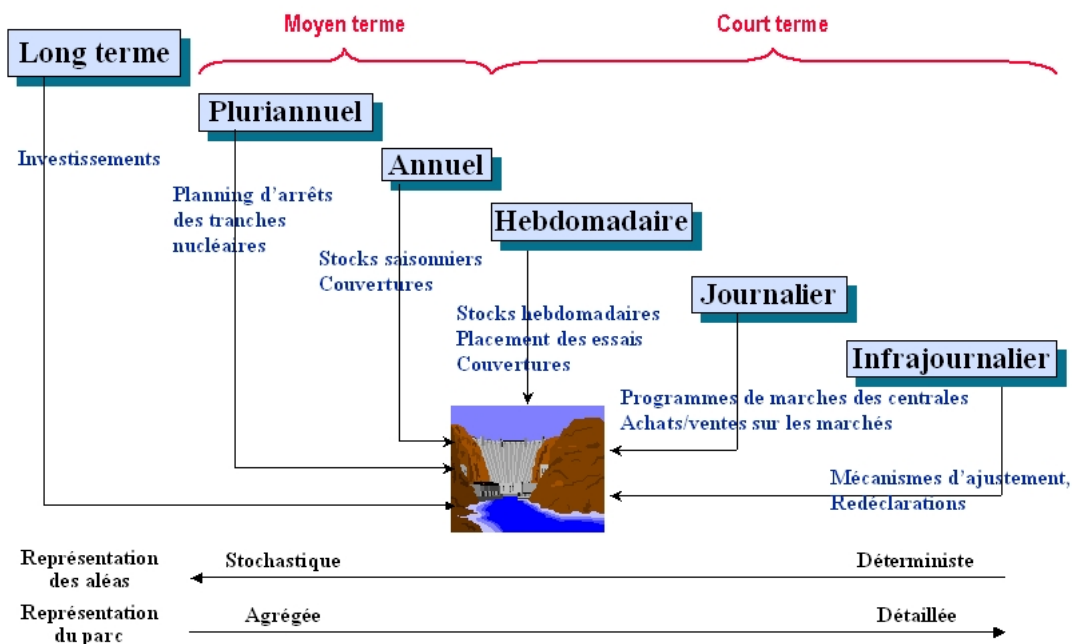


FIG. 1.1 – Horizons de temps considérés en gestion de production et décisions correspondantes

Les contraintes doivent être d'autant plus détaillées que l'on est proche du temps réel. Autant nous pouvons simplifier la modélisation du parc pour des horizons comme l'annuel (en agrégeant les stocks hydrauliques ou en supprimant les contraintes techniques dynamiques par exemple) puisque seuls des indicateurs sont calculés, autant la modélisation doit être fine aux horizons court-terme où le planning final, prenant en compte toutes les contraintes techniques, doit être déterminé. Cette différence de modélisation est une particularité de ce découpage temporel et a pour principale conséquence de rendre les modèles du court-terme assez complexes puisque très gros (à cause de la grande taille du parc de production), très détaillés (introduisant des non-linéarités et de la non-convexité) et donc très contraints.

ainsi une utilisation transparente du réseau.

L'ouverture du secteur électrique européen, avec la mise en place de marchés de l'électricité, a de nombreuses incidences sur la gestion de la production. Les marchés constituent pour chaque producteur un nouvel outil de couverture de ses risques physiques et financiers, mais également une source supplémentaire pour dégager des gains potentiels (vente de surplus de production, opportunités d'arbitrage). Il est donc intéressant d'introduire les décisions d'achats et ventes au sein même des modèles de planification. Tout comme il devient aussi impératif de maîtriser le risque lié à l'incertitude sur des paramètres, autrefois bien connus comme la demande de consommation, et à la volatilité de nouvelles variables telles que les prix de marché.

La libéralisation a aussi conduit les producteurs à revoir leur processus décisionnel et donc l'ensemble des flux d'indicateurs. Cette adaptation a amené EDF à évaluer les moyens de mieux gérer la production journalière (grâce notamment à la transmission d'indicateurs économiques élaborés sur un horizon plus long), et à définir ses stratégies de prises de position sur les marchés. C'est ainsi qu'a émergé peu à peu la nécessité d'une gestion prévisionnelle de la production à l'horizon hebdomadaire/mensuel, échelle de temps qui était peu étudiée auparavant.

1.1.4 Emergence d'une gestion hebdomadaire/mensuelle de la production

Le rôle actuel de la gestion hebdomadaire consiste à donner une vision du futur à l'horizon journalier en fournissant des indicateurs économiques d'arbitrage et de gestion des stocks hebdomadaires, en particulier les options tarifaires et l'hydraulique. Ce rôle de "sentinelle du journalier" est d'une grande importance dans la chaîne de décision de production d'électricité. En anticipant ainsi l'état du parc EDF sur les deux semaines à venir, nous permettons aux équipes en charge des gestions journalière et infra-journalière d'optimiser au mieux les plannings de production, compte tenu d'une certaine vision du futur.

Le second rôle de la gestion hebdomadaire, dont la présente thèse constitue une partie des études exploratoires, consiste d'une part à gérer la marge de production, définie comme l'écart entre l'offre totale et la demande totale en électricité, compte tenu des aléas affectant le système électrique, et d'autre part, à maîtriser les risques associés à la gestion de cette marge. Il s'agit d'une problématique conjointe de couverture et de gestion active de la marge. Si la demande ne peut être satisfaite, c'est le black-out ou défaillance physique. L'objectif est donc d'assurer la non-défaillance avec une bonne probabilité. Pour cela, la marge d'exploitation est comparée à une marge de sécurité pour que la probabilité d'avoir une marge d'exploitation négative soit égale à un niveau de consigne.

Lorsque le risque de défaillance est trop élevé, nous cherchons à nous couvrir de ce risque par des décisions de couverture telles que :

- le changement de la structure du parc de production : nous différons typiquement un départ en maintenance ou en essai d'une ou plusieurs centrales ;
- le déclenchement d'options tarifaires (contrats d'effacement) ;
- l'achat sur les marchés.

La problématique de couverture physique consiste donc à déterminer quelles décisions, optimales selon un certain critère économique, il faut prendre pour se couvrir contre le risque de défaillance. Ce risque, le “risque $p\%$ ”, est un critère qui consiste à imposer que l'écart entre l'offre et la demande (la marge) soit supérieur à une valeur de sécurité sur l'ensemble des scénarios considérés dans au moins $100 - p\%$ des cas. Ce critère n'est en réalité qu'une formulation mathématique des dispositions réglementaires auxquelles un producteur d'électricité doit faire face. Ceci nécessite de savoir évaluer la marge, compte tenu des aléas affectant le système électrique, d'évaluer le risque de défaillance, de définir une fonction économique d'arbitrage des décisions de couverture et de résoudre le problème en résultant, tenant compte des contraintes.

Or actuellement, si EDF dispose d'un outil de gestion hebdomadaire permettant de satisfaire le rôle actuel de “préparation de la gestion journalière”, il n'existe pas d'outil opérationnel pour satisfaire les besoins relatifs au rôle de gestion de la marge et du risque. Cet état de fait ne permet donc pas, actuellement, d'aborder les problématiques de mesure du risque de défaillance, de la couverture, voire celle du trading, en coordination avec la production à l'horizon hebdomadaire.

1.1.5 Sujet de thèse

Cette thèse doit permettre d'apporter des réponses à la problématique de la gestion des risques physiques dans chacun des rôles de l'horizon hebdomadaire. Pour le premier rôle, il s'agit principalement d'explorer de nouvelles méthodologies qui pourraient permettre d'intégrer de manière satisfaisante les aléas dans le problème de la gestion de la production. Pour le second rôle, nous ne disposons d'aucun existant et tout est à faire. Il faut notamment apporter des éléments de réponse sur la modélisation de cette problématique, sur les aléas à prendre en compte, et évidemment sur la résolution du problème ainsi obtenu.

Cette thématique est de base très complexe et d'un enjeu majeur, ce qui donne de l'importance à chaque élément de réponse apporté.

1.2 Contexte : méthodes d'optimisation utilisées

Dans cette thèse, nous avons essentiellement appliqué deux méthodes d'optimisation dans l'incertain : l'optimisation robuste et la programmation stochastique avec règles de décision constantes par morceaux. Nous allons dans cette section les présenter brièvement. Nous retrouverons une présentation plus détaillée de chacune dans les chapitres qui leur sont dédiés.

1.2.1 L'optimisation robuste

L'*optimisation robuste* traite de problèmes d'optimisation dans lesquels les coefficients des contraintes et de l'objectif sont incertains. Son objectif est de donner une solution opérationnelle à ces problèmes en des temps raisonnables contrairement aux approches plus traditionnelles, basées sur un modèle probabiliste, qui se heurtent à des difficultés numériques considérables. En effet, il est en général très difficile :

- d’avoir un modèle probabiliste de chaque aléa qui soit pertinent ;
- de calculer des espérances mathématiques avec des distributions multidimensionnelles avec une précision raisonnable.

Et même si ces deux difficultés sont surmontées, le problème d’optimisation est alors d’une telle complexité qu’il est impossible à résoudre, en particulier lorsqu’il s’agit de problèmes dynamiques avec plus de deux périodes.

L’optimisation robuste propose une approche de protection contre le “pire des cas”, pire des cas à prendre dans un *ensemble d’incertitude* à déterminer. La construction de cet ensemble des possibles ne nécessite pas d’hypothèses probabilistes. Le critère d’espérance étant remplacé par la notion de “pire des cas”, le problème du calcul des espérances est évacué et conduit à un problème d’optimisation de même nature et de taille modérément supérieure. L’idée principale consiste à choisir un ensemble d’incertitude ne contenant que les réalisations les plus réalistes, excluant donc des réalisations pouvant conduire à une protection excessive et à un coût de protection exorbitant. Cette formulation par le pire des cas conduit à remplacer la satisfaction d’une contrainte déterministe, par la satisfaction d’un nombre infini de contraintes, chacune de ces contraintes correspondant à une réalisation particulière de l’aléa dans l’ensemble d’incertitude. Le problème d’optimisation qui en résulte appartient à la programmation mathématique semi-infinie dont la complexité semble bien éloignée de l’objectif de simplicité et de facilité de calculs évoquée plus haut. Ce n’est pourtant pas le cas, car sous des hypothèses appropriées sur les ensembles d’incertitude (par exemple, des polyèdres ou des ellipsoïdes) et sur les contraintes initiales (par exemple la linéarité), il s’avère possible de remplacer cet ensemble infini de contraintes par un nombre fini d’autres contraintes. L’outil qui permet cette transformation est un théorème de dualité forte. Nous obtenons grâce à lui une formulation explicite de l’équivalent robuste comme un problème de programmation convexe de dimension raisonnable que nous pouvons résoudre par des algorithmes efficaces.

De nombreuses extensions de l’optimisation robuste se sont développées dans la littérature. En particulier, nous pouvons citer l’extension du concept de programmation robuste aux problèmes à plusieurs périodes. Cependant, elle pose un problème conceptuel sérieux : comment traduire de manière satisfaisante ce que nous appelons improprement le caractère non-anticipatif des décisions, qui n’est rien d’autre que l’expression d’ignorance du futur au moment de la prise de décision. Dans [15], le concept de règles de décision affinement adaptatives est introduit. Grâce à cet artifice, les commandes robustes sont plus adaptatives et non anticipatives. Le choix de la règle est pleinement empirique : nous pouvons choisir par exemple de ne considérer que le dernier aléa réalisé ou tout l’historique. L’approche robuste multi-périodique doit donc s’interpréter comme une heuristique permettant l’obtention d’une politique sous-optimale. Cependant, comme aucune approche n’est, à l’heure actuelle, satisfaisante pour résoudre de tels problèmes, l’approche robuste avec règles de décision mérite d’être étudiée.

Durant cette thèse, les problèmes que nous avons considérés se sont tous avérés être linéaires. Sous cette hypothèse de linéarité du problème originel, nous pouvons montrer que la *contre-partie robuste* obtenue est elle-aussi formulée comme un problème d’optimisation linéaire. Cette contre-partie robuste introduit des variables et des contraintes auxiliaires, mais en nombre suffisamment réduit pour que le problème transformé soit traitable par les solveurs actuellement disponibles sur le marché, que ce soient des logiciels libres ou des produits commerciaux. L’introduction de règles affines en l’aléa ne remet pas en cause la linéarité de l’équivalent robuste. Cependant, la

taille du problème devient encore plus importante.

D'autre part, l'approche par optimisation robuste semble orthogonale à une vision classique basée sur le calcul des probabilités. Néanmoins, il existe une parenté évidente entre la satisfaction d'une contrainte pour le pire des cas choisi dans un ensemble d'incertitude et l'optimisation sous contrainte en probabilité [27, 28, 25] connue sous le nom de *Chance-constrained programming*. Nous pouvons même affirmer que l'optimisation robuste est une façon d'approximer les contraintes en probabilité par des contraintes convexes. Le niveau de probabilité étant assuré par la taille de l'ensemble d'incertitude. De nombreux théorèmes sont d'ailleurs apparus récemment pour borner la probabilité de non satisfaction d'une contrainte par la solution robuste.

La première formulation d'une solution robuste à un problème de programmation mathématique avec coefficients incertains peut être attribuée à Soyster [58] au début des années 70. Le domaine a pris son essor beaucoup plus tard, dans la fin des années 90, grâce au travail de pionniers de Ben-Tal et Nemirovski [16], mais également de El-Ghaoui et Lebret [33]. Depuis, l'optimisation robuste est devenue un des domaines de recherche en optimisation des plus actifs. Les principaux développements théoriques se trouvent dans les articles de Ben-Tal et co-auteurs [17, 15, 13, 19] ou Bertsimas [21, 20]. Le nombre d'applications et la variété des domaines qu'ils couvrent ne cessent de croître. Nous pouvons citer par exemple la finance et les mesures de risque [23, 24, 26, 36], la gestion des stocks et des chaînes d'approvisionnement [1, 14, 22] ou les télécommunications [49]. Cette approche peut également s'appliquer en complément de techniques d'optimisation comme les contraintes en probabilités [25], la programmation dynamique [44], ou la programmation stochastique [29].

1.2.2 La programmation stochastique avec règles de décision constantes par morceaux

Une formulation anticipative dite en boucle ouverte n'utilise pas la connaissance des aléas réalisés dans le passé. En effet, les décisions sont prises à l'instant initial t_0 sur toute la période d'étude. Or, en un certain pas de temps t , le décideur a la connaissance des réalisations des aléas sur la période $[t_0; t - 1]$. Il est donc bien plus réaliste de faire dépendre les commandes de l'observation des aléas réalisés. Cette dépendance a aussi pour effet de satisfaire la contrainte de non anticipativité : le décideur ne doit pas connaître parfaitement le futur pour optimiser ses décisions. En effet, si en un certain pas de temps t^* tout le futur est dévoilé, le problème devient un problème à 2 étapes. Dans ce cas où les commandes dépendent du temps et de l'histoire des processus stochastiques, nous disons que les variables sont en boucle fermée.

L'approche classique dans la littérature pour résoudre des problèmes à plusieurs étapes est la programmation stochastique sur arbre de scénarios. La première étape de cette approche est la construction de l'arbre : après une analyse détaillée des processus stochastiques, nous devons être capables de générer un grand peigne de scénarios. Ces scénarios sont agrégés en supposant que les décisions en un instant doivent être identiques pour tous les scénarios les plus proches au sens d'une certaine distance. Une fois l'arbre construit, nous pouvons formuler l'équivalent déterministe en attribuant une commande à chaque nœud de l'arbre, c'est à dire une commande à chaque réalisation de l'aléa. Si le problème stochastique original est linéaire, l'équivalent déterministe obtenu sera toujours un problème d'optimisation linéaire mais d'une taille bien plus

importante. La difficulté majeure de cette approche est sa complexité : si nous souhaitons modéliser de façon cohérente l'aléa, le nombre de branchements doit croître exponentiellement dans le temps (contrainte de raffinement). D'ailleurs, dans [56, 54, 57, 55], les auteurs avancent que "les problèmes multi-étapes, même linéaires, avec des recours complets sont généralement numériquement intraitables même pour une précision faible". Ce constat n'empêche pas que "certains problèmes puissent être résolus efficacement". Dyer et Stougie avancent pour leur part que "la programmation stochastique multi-étapes est au moins aussi complexe que les problèmes combinatoires difficiles". Pour remédier à cela, il existe des techniques de réduction d'arbre [41, 32] qui permettent de réduire la taille de l'arbre tout en gardant l'information principale. Utiliser la programmation stochastique oblige donc à faire un arbitrage entre les temps de calculs et la précision de la solution.

C'est pour tenter de remédier à ces problèmes qu'a été introduite la *Programmation Stochastique avec Règles de Décision Constantes par Morceaux* (Stochastic Programming with Step Decision Rules (SPSDR)). Cette heuristique a été initiée dans l'article [59] par J. Thénier et J.-P. Vial. Elle a introduit sur des problèmes de contrôle sans contrainte sur les états et avec un aléa unidimensionnel.

Dans SPSDR, le peigne de scénarios initial est conservé et, à chaque instant, nous effectuons des partitions \mathcal{P}_t de scénarios (par la même technique que pour la détermination des nœuds d'un arbre en programmation stochastique classique). Chaque paquet de scénarios obtenu est appelé un *ensemble de décision*. Le principe de cette heuristique est relativement simple et basé sur deux hypothèses :

1. puisque la contrainte de raffinement est la cause de l'accroissement exponentiel du nombre de nœuds de l'arbre en programmation stochastique, nous décidons de la supprimer. Le nombre d'éléments de \mathcal{P}_{t+1} peut être plus faible que celui de \mathcal{P}_t (cf. figure (6.3)).

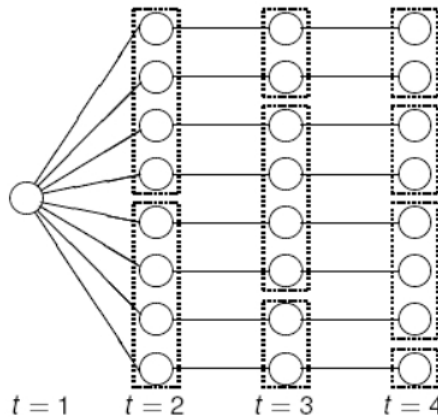


FIG. 1.2 – Construction des ensembles de décision de l'approche SPSDR

2. nous imposons une unique décision par ensemble de décision. Une règle de décision s'adaptant à la réalisation de l'aléa est ainsi construite.

Nous pouvons détailler les idées principales de l'heuristique :

- **Construction des ensembles de décision** : les ensembles de décision sont construits en partitionnant chaque sous-scénario en p paquets, chaque paquet regroupant les sous-scénarios par proximité au sens d'une certaine distance. Un sous-scénario \hat{s}_t est la restriction du scénario s_t sur la période $1, \dots, t$. Ce problème de répartition des sous-scénarios dans p paquets est en fait un problème de p -médianes. Ce problème est un problème classique en Recherche Opérationnelle et est connu pour être NP-difficile [45]. Des heuristiques peuvent être utilisées pour résoudre ce problème ;
- **Construction des règles de décision** : Nous venons de construire des ensembles de décision I_t^k , $k = 1, \dots, K_t$ à chaque période de temps t . Sur chaque ensemble de décision nous imposons une décision unique v_t^k , déterminée après optimisation sur un pool de scénarios de référence. Si nous souhaitons reconstruire la règle de décision sur un scénario réalisé σ , il suffit de trouver de quel ensemble de décision ce scénario appartient à chaque instant pour lui appliquer la décision correspondante. Introduisons la fonction indicatrice $\mathbb{1}_{\{\sigma_t \in I_t^k\}}$ qui vaut 1 si le sous-scénario $\hat{\sigma}_t$ est associé à l'ensemble de décision I_t^k et 0 sinon. Comme une seule décision ne peut être prise sur chaque scénario pour un pas de temps, nous pouvons récrire la règle de décision sous la forme :

$$u_t(\sigma_t) = \sum_{k=1}^{|I_t|} \mathbb{1}_{\{\sigma_t \in I_t^k\}} v_t^k, \quad \forall t = 2, \dots, T-1. \quad (1.1)$$

Les variables v , qui sont les décisions réelles, sont donc, en tant que somme de fonctions indicatrices, des *règles de décision continues par morceaux*.

- **Utilisation de K experts** : résoudre le problème initial sur un grand nombre de scénarios peut s'avérer très coûteux en temps de calculs. Pour remédier à ce problème, nous pouvons utiliser l'avis d'experts indépendants comme dans [47]. Au lieu de résoudre un gros problème d'optimisation, nous proposons de résoudre K fois le problème sur des échantillons réduits (tirés aléatoirement dans le pool initial) de scénarios. Ainsi, nous ne reconstruisons plus une seule règle $u(\sigma)$ pour le scénario σ mais désormais K règles $u^l(\sigma)$. La commande finale appliquée sera une combinaison convexe des règles obtenues par chaque expert :

$$u_t(\sigma_t) = \sum_{l=1}^K \alpha_l u_t^l(\sigma_t), \quad \sum_{l=1}^K \alpha_l = 1, \quad \forall t = 1, \dots, T.$$

- **Validation** : généralement, en programmation stochastique, la qualité de la solution obtenue n'est jamais évaluée. Dans SPSDR, la phase de validation est importante et permet de valider la pertinence de la solution proposée. Pour ce faire, nous appliquerons sur chaque scénario du pool de référence la règle de décision obtenue et nous évaluerons des indicateurs classiques comme les coûts de gestion, la distribution des coûts,...

Cette approche a été testée dans [59] sur un problème de chaîne d'approvisionnement. Les expériences numériques sur un jeu d'essais particulier montrent que SPSDR permet d'obtenir une erreur de prédiction (écart entre le coût optimal obtenu en optimisation et le coût obtenu

en simulation) moins importante que pour la programmation stochastique classique sur arbre (PSA) : 1% d'erreur contre 12%. L'amélioration du coût moyen en simulation de SPSDR par rapport à PSA est de l'ordre de 5%. Enfin, l'optimisation robuste permet d'obtenir des stratégies moins risquées mais plus coûteuses (une augmentation de coût de l'ordre de 10 à 20% selon le niveau de couverture).

1.3 Synthèse des résultats

Dans le cadre de cette thèse, les enjeux industriels sont particulièrement importants. L'attente est forte sur la capacité à proposer des approches (ou des pistes de réflexion) opérationnelles sur des problèmes industriels concrets. Nous entendons par là des approches qui apportent des réponses validées expérimentalement, comparées à des approches existantes (dans l'éventualité où c'est possible) et dans des temps de calculs qui soient maîtrisés. Nous avons donc porté nos efforts dans cette direction. Les problèmes considérés sont la gestion d'une vallée hydraulique et la gestion du risque de défaillance physique. Ils sont tous les deux d'un enjeu important, liés au risque de défaillance physique et permettent de contribuer aux deux rôles de la gestion hebdomadaire. Ces problèmes industriels sont des problèmes complexes, où aucune approche n'est entièrement satisfaisante. Nous avons ainsi tiré des idées d'autres domaines pour les appliquer à la problématique de la gestion de production électrique.

Par ce biais, la contribution industrielle est double et constituée :

- d'un apport méthodologique ;
- d'expériences numériques prometteuses.

Nous pouvons présenter désormais de façon plus détaillée les travaux inhérents à cette thèse.

1.3.1 La gestion robuste d'une vallée hydraulique

Motivations

La gestion des stocks hydrauliques est un enjeu capital dans le domaine de la gestion de la production en électricité. Les particularités de l'hydraulique sont : un coût de production nul et un temps d'accessibilité faible. Elle constitue ainsi une ressource de substitution aux unités de production les plus chères. Il est en effet bien plus avantageux de conserver son stock d'eau pour les pointes de consommation, moments où les unités chères comme les centrales à fioul sont utilisées. Dans le cadre du découpage temporel, la production hydro-électrique est gérée sur différents horizons de temps. A l'annuel, nous lui déterminons une valeur, appelée la valeur de l'eau, fonction de la saison dans laquelle nous nous trouvons, de son niveau de volume et du niveau des stocks complémentaires. Ces valeurs d'usage étant connues, nous pouvons ainsi à court-terme prendre la meilleure politique de turbinage/pompage en arbitrant entre l'utilisation immédiate de l'eau (rémunérée au prix de l'unité de production la plus chère du moment, soit le coût marginal du système) et la conservation du niveau d'eau (rémunérée à la valeur de l'eau). Cependant, cette optimisation du planning de production n'est pas évidente en raison des diverses contraintes techniques (paliers de production, rendements continus par morceaux,...) et

de l'aléa d'hydraulicité (pluie, fonte des neiges,...) affectant chaque réservoir. D'ailleurs, ces apports ont un rôle important dans la gestion hydraulique. Leur variabilité est assez élevée et il est assez difficile de les prévoir (car ils sont issus de phénomènes physique et géologique, difficiles à modéliser). Or, une grande quantité d'apports non anticipée peut entraîner des turbinages d'urgence afin que les niveaux d'eau ne dépassent pas des cotes d'alerte, tout comme un manque d'apports peut inciter un blocage de la production pour assurer un débit minimum à la vallée. De telles contraintes sont d'autant plus importantes à satisfaire qu'elles font partie des obligations réglementaires d'EDF vis à vis des collectivités locales.

Une gestion hydraulique capable d'être optimale à l'horizon hebdomadaire, appréhendant l'aléa d'hydraulicité et satisfaisant au mieux les contraintes techniques est un véritable enjeu. Cependant, ce problème revient à établir une stratégie de turbinage à un pas de temps initial sur l'ensemble d'une période avec des réalisations d'hydraulicité périodiques. Nous sommes donc face à un problème stochastique à plusieurs périodes très difficile. A l'heure actuelle, la théorie de l'optimisation dans l'incertain ne donne pas de réponses pleinement satisfaisantes pour de tels problèmes. Dans les pays à forte capacité hydraulique (Norvège, Brésil,...), l'approche la plus utilisée est la programmation dynamique stochastique duale [50, 46, 53]. Cette approche leur permet de résoudre un problème avec un nombre de stocks supérieur à ce que la programmation dynamique stochastique standard leur permettrait de résoudre (autour de la dizaine). D'autre part, les stocks hydrauliques étant de grande taille, les unités de type thermique sont peu nombreuses ce qui permet de résoudre le problème frontalement. Unger [60] propose également une approche où les unités de production sont gérées face au marché, c'est à dire avec une production dépendante des prix de marché. Une approche par programmation stochastique où l'aléa d'hydraulicité est représenté par un arbre a été étudiée par Römisch et al [48, 42, 38, 37, 30], Frauendorfer et Güssow [35, 39] ou Fleten et Kristoffersen [34]. Le problème de la programmation stochastique est que, pour représenter l'aléa de manière satisfaisante, la taille de l'arbre croît exponentiellement dans le temps. Or, l'horizon hebdomadaire comportant 7 à 12 jours avec un pas de temps bi-horaire, il semble peu cohérent d'appliquer une telle approche. Prékopa a introduit une approche basée sur l'utilisation de contraintes en probabilité [51]. Cette approche cherche à satisfaire les contraintes de bornes sur les volumes avec une certaine probabilité. En théorie, cette approche correspond à nos besoins. Mais, pour transformer les contraintes en probabilité en contraintes linéaires, il est nécessaire de calculer des quantiles de la distribution des apports. Il est donc nécessaire de connaître la distribution de probabilité des apports hydrauliques. En pratique, cette distribution est inconnue ce qui rend cette approche difficile à mettre en place dans notre cas. Or, il existe une approximation convexe des contraintes en probabilité ne nécessitant aucune connaissance sur la distribution probabiliste de l'aléa : l'optimisation robuste. La résolution du problème hydraulique par optimisation robuste, qui se justifie complètement sur un point de vue opérationnel, est une approche originale dans le cadre de la planification électrique.

Application de l'optimisation robuste à la gestion d'une vallée hydraulique

Nous appliquons dans cette partie de thèse l'optimisation robuste présentée en 1.2.1.

Une vallée hydraulique est une succession de réservoirs et d'usines, chacune constituée de plu-

1.3. SYNTHÈSE DES RÉSULTATS

sieurs groupes de turbinage et/ou pompage. En pratique, une vallée est sujette à de nombreuses contraintes :

- contraintes de bornes sur les turbines/pompages de chaque groupe ;
- contraintes de bornes sur les volumes des réservoirs ;
- paliers de production ;
- non changement de sens et gradient de débit ;

La gestion d'une vallée doit tenir compte de ces contraintes, des équations de conservation de masse qui régissent la dynamique du système et enfin du coût marginal de l'électricité et de la valeur finale d'usage de l'eau. Dans la pratique, l'indisponibilité des centrales, planifiée ou subie, est un facteur important intervenant dans la gestion.

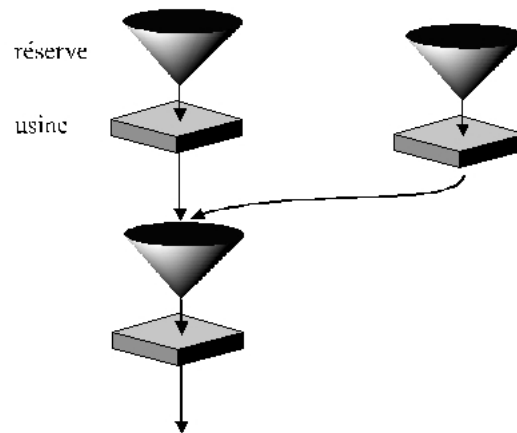


FIG. 1.3 – Représentation d'une vallée hydraulique

Malgré la présence de non-linéarités dans une modélisation fine de la gestion du parc hydraulique, une représentation par programmation linéaire d'une version déterministe du problème est parfaitement acceptable. La fonction objectif à maximiser est un arbitrage à chaque instant entre la conservation de l'eau rémunérée à la valeur de l'eau et la production d'électricité rémunérée, elle, au coût marginal du système (c-à-d le coût de l'unité de production la moins chère à utiliser dans le cas d'une unité de demande supplémentaire). Les commandes de turbinage optimales, tout en restant dans les plages de fonctionnement des usines, doivent contribuer à équilibrer le bilan d'eau de chaque réservoir. La satisfaction des bornes sur les volumes d'eau est la contrainte la plus forte car sujette à la réglementation portant sur les niveaux de volume.

Le problème de gestion d'une vallée hydraulique à l'horizon hebdomadaire (sur la période 1 à H) peut ainsi s'écrire :

$$\min_T \quad \sum_g \rho_g \sum_{\tau \leq T} c_\tau T_{g\tau} + \sum_r \omega_r (V_r^0 - V_r^H) \quad (1.2)$$

$$V_r^t = V_r^0 + A_r^t + \sum_{\tau \leq t} \sum_g \delta_{gr} T_{g\tau}, \quad \forall r, \quad \forall t \leq H \quad (1.3)$$

$$\underline{V}_r^t \leq V_r^t \leq \bar{V}_r^t, \quad \forall r, \quad \forall t \leq H \quad (1.4)$$

$$\underline{T}_g^t \leq T_{gt} \leq \bar{T}_g^t, \quad \forall g, \quad \forall t \leq H. \quad (1.5)$$

où les T_{gt} sont les quantités d'eau lâchées ou pompées par le groupe g au pas de temps t (bornées par les quantités \underline{T}_g^t et \bar{T}_g^t), V_r^t est le volume du réservoir r au pas de temps t (borné par les quantités \underline{V}_r^t et \bar{V}_r^t), δ_{gr} décrit la topologie de la vallée (paramètre valant 1 si le groupe g est en amont du réservoir r , -1 s'il est en aval de r et 0 s'il n'est pas connecté à r), ρ_g le rendement du groupe g , c_t le coût marginal du système en t , ω_r la valeur de l'eau du réservoir r et A_r^t la quantité d'apports naturels reçue au réservoir r et au pas de temps t .

Un concept clé de l'optimisation robuste est l'ensemble d'incertitude. L'ensemble d'incertitude contenant les aléas sur les apports d'eau devraient être décrits par des boules obtenues à l'aide de la norme ℓ_2 . Le pire des cas avec cette norme est en effet plus réaliste que le pire des cas obtenu à l'aide de normes usuelles : à titre d'exemple, le pire des cas avec ℓ_∞ représente la situation où tous les aléas prennent leur pire valeur à chaque instant, situation trop contraignante. D'autre part, en transformant les apports en relation affines dépendantes d'un bruit blanc, la probabilité de présence dans la boule ℓ_2 est donnée à l'aide d'une loi de Chi-deux. Nous pouvons ainsi calibrer la taille de l'ensemble d'incertitude aisément. Cependant, l'utilisation d'une telle norme conduit à un problème d'optimisation conique quadratique. Pour conserver un problème d'optimisation linéaire, nous faisons le choix d'approximer la boule ℓ_2 par un polyèdre obtenu par l'intersection d'une norme ℓ_1 et d'une norme ℓ_∞ .

Si ces ensembles d'incertitude sont des polyèdres, la version robuste du problème est encore un programme linéaire, de taille légèrement plus grande, et donc, dans certains cas, de complexité voisine de celle du modèle déterministe (en terme de taille et donc de temps de résolution). Cette modélisation a cependant l'inconvénient majeur de ne pas prendre en compte les possibilités d'ajustement (ou de recours) attachées aux périodes futures. Pour remédier à cette limitation, les auteurs de [15] ont utilisé le concept de règle de décision linéaire. Il s'agit du second concept clé nécessaire pour traiter par optimisation robuste les problèmes dynamiques. L'utilisation de règles de décision linéaires conduit de nouveau à la formulation du problème stochastique global comme un programme linéaire, certes de plus grande taille, mais encore gérable par les outils de modélisation et de résolution actuellement disponibles. Dans le cas de l'hydraulique, les règles de décision prennent la forme de fonctions linéaires des observations journalières sur les apports d'eau.

Cette approche robuste a été testée sur une vallée de 3 réservoirs et 6 groupes, gérée sur une période de 7 jours (84 pas de temps). Cette vallée, de taille raisonnable, est représentative d'une partie des vallées hydrauliques françaises. Il n'existe pas de solution reconnue optimale à laquelle nous pouvons comparer la solution robuste. Pour valider notre approche, nous avons donc procédé à des simulations. Cette phase de validation est généralement peu mise en avant

dans la littérature. Cependant, elle semble incontournable si nous voulons évaluer la qualité de la solution. Pour cela, nous avons tiré 10 000 scénarios et avons mis en œuvre deux approches différentes :

- Dans l’approche *robuste à révisions périodiques* (RRP), seules les contraintes de la première journée sont rendues robustes. Pour les journées suivantes, les contraintes ne s’appliquent que pour les prévisions d’apports d’eau. La révision périodique consiste à lancer une optimisation au début de chaque journée, en tenant compte des réalisations de l’aléa tout au long de la journée, et donc de niveaux de volume réactualisés. La révision périodique introduit une capacité d’adaptation. La prise en compte dans l’optimisation d’un horizon dépassant le premier jour—même sur un mode déterministe— donne une cohérence aux décisions du premier jour.
- Dans l’approche *robuste à ajustements linéaires* (RAL), des règles de décision pour les jours 2 et suivants introduisent directement dans le problème d’optimisation la faculté d’ajustement aux aléas manifestés. Le modèle résultant est plus complexe, mais il est théoriquement implémentable sur tout l’horizon d’étude sans réactualisation.

Comme élément de comparaison, nous avons retenu tout d’abord une politique à information parfaite (IP). Dans cette politique, nous effectuons une optimisation sur chaque scénario, comme si la chronique des apports d’eau était parfaitement connue dès le début de l’horizon. Cette politique est une référence idéale, non réalisable. Elle n’entraîne aucune violation de contrainte et sa performance économique est une borne inférieure absolue pour toute autre politique. Le deuxième élément de comparaison est une politique que nous nommons *déterministe à révision périodique* (DRP). Cette politique peut être considérée comme assez proche de celle suivie en pratique en exploitation : à chaque période, nous optimisons sur une version déterministe du problème, où les apports d’eau sont fixés à leur prévision. Des heuristiques ont même été intégrées pour éviter des comportements impossibles en pratique (continuer à turbiner sous \underline{V}_r^t , stocker au dessus de \bar{V}_r^t). Cette politique possède ainsi une certaine flexibilité qui la rend un peu plus comparable² qu’une *politique déterministe fixe* (DF).

Les expériences numériques montrent que les approches robustes assurent une excellente satisfaction des contraintes. Cela est particulièrement vrai pour RRP, mais, de manière surprenante, aussi pour RAL où toutes les décisions sont fixées (sous forme de règles) à l’origine de la période de planification. Le surcoût par rapport à IP n’est que de 0,57 % et 1,15 %, pour respectivement RRP et RAL. Par contre, DRP est incapable de satisfaire les contraintes. En particulier, nous observons des déversés sur pratiquement tous les scénarios et la valeur moyenne de ces déversés est si importante que la performance économique sur les simulations ne peut être prise en considération³.

Les approches robustes étudiées offrent une alternative crédible pour la prise en compte opérationnelle d’aléas dans le problème de gestion de la production hydraulique. De plus, nous pouvons imaginer appliquer l’approche à ajustements linéaires à des problèmes de décisions fermes dans l’incertain sur une grande période, comme peut l’être un problème de couverture financière.

²Une politique déterministe fixe (DF) où les décisions de période 2 et suivantes seraient fixées rigidement dès la première période, donc sans procéder à des réoptimisations ultérieures, est totalement inopérante, et viole massivement les contraintes.

³En effet, nous observons que les violations importantes des contraintes permettent d’améliorer le coût de IP !

Cette approche peut aussi être améliorée par une révision périodique. Nous n'avons pas étudié cette possibilité faute d'outils plus performants que les logiciels de modélisation et de résolution disponibles en « logiciel libre » utilisés pour cette étude de faisabilité.

Une maquette en Matlab a été implémentée et permet de tester toutes les approches proposées sur diverses vallées hydrauliques. Les phases d'optimisation sont résolues à l'aide du logiciel libre de programmation linéaire GLPK ⁴.

Pistes d'améliorations

La principale amélioration pouvant être apportée à cette approche est l'introduction d'un aléa de prix dans la fonction objectif. L'aléa est ainsi en produit avec une variable de décision du problème. Dans le cas RRP, la prise en compte robuste ne conduit à aucune difficulté supplémentaire. Par contre, l'application d'ajustements linéaires (modèles RAL) conduit à des produits d'aléas. L'obtention du pire des cas est alors une question ouverte. Cependant, des approches comme la programmation conique [18] pourraient résoudre ces problèmes. Un saut apparent dans la difficulté est introduit, mais, fondamentalement, ces approches ne sont pas plus complexes que la programmation linéaire.

Le modèle robuste peut également être amélioré par deux biais supplémentaires. L'introduction de règles de décision avec paramètres de recours fixes (section 3.8.2) permet d'appliquer des stratégies différentes selon que l'aléa est plus ou moins élevé par rapport à la moyenne. Nous pouvons aussi imaginer appliquer la violation contrôlée des contraintes (section 3.9). Des violations peuvent être ainsi autorisées mais maîtrisées à l'aide de pénalités. Ces approches pourraient enrichir le modèle robuste et surtout le rendre encore plus réaliste dans le contrôle des volumes.

Enfin, l'expérimentation de ces approches robustes dans le modèle hebdomadaire existant paraît être une étape nécessaire pour l'évaluation de leur performance. En particulier l'approche RRP dont l'accroissement de la taille du modèle devrait être tout à fait raisonnable et ainsi donner lieu à des temps de calculs du même ordre que le modèle actuel.

1.3.2 La gestion du risque physique

Motivations

Dans la deuxième partie de cette thèse, nous nous intéressons au second rôle de l'horizon hebdomadaire. Nous cherchons ainsi à optimiser une couverture financière en vue d'assurer une marge de production positive, sur 100-p % des scénarios possibles (obligation de service public), la marge étant définie comme la différence entre l'offre et la demande.

Pour cela, il est nécessaire de disposer de méthodes et d'outils pour :

- évaluer la capacité maximale de production thermique et hydraulique (dont la somme représente l'offre de production) tenant compte des aléas affectant le système électrique ;

⁴GNU Linear Programming Kit (GLPK) est un kit d'optimisation libre contenant un outil de modélisation, sous-ensemble d'AMPL, et un solveur, glpsol, permettant de résoudre des problèmes linéaires à variables mixtes.

- résoudre le problème d’optimisation de minimisation du coût de la couverture, constituée d’actifs financiers et potentiellement physiques.

Actuellement, ces besoins ne sont satisfaits que par des heuristiques “métier”, sous-optimales. Dans la perspective du développement de telles méthodologies à l’horizon hebdomadaire (voire mensuel), nous avons étudié la possibilité de créer un outil global de gestion (production et risque). Un tel problème est très complexe car stochastique avec de nombreux aléas, de très grande taille et non convexe. Pour faire face à ce problème, impossible à résoudre numériquement, nous devons trouver des simplifications. C’est ainsi que nous proposons une approche en deux étapes. La première étape consiste à évaluer de façon probabiliste la marge de production et le risque $p\%$, compte tenu des aléas affectant le système électrique. La seconde étape est consacrée à la formulation mathématique du problème d’optimisation sous contraintes en probabilité et à sa résolution.

L’amélioration du calcul de marge

La marge est l’écart entre la capacité maximale de production du parc EDF (thermique et hydraulique) et la demande. Il s’agit du processus stochastique à la base du problème de gestion des risques physiques à l’hebdomadaire. Si nous ne sommes pas capables de l’estimer convenablement, tous les efforts pour écrire un modèle d’optimisation pertinent peuvent s’avérer vains. Actuellement, la marge est calculée (pour quelques scénarios) en supposant que tous les aléas sont gaussiens et en sommant les puissances maximales des centrales (thermique et hydraulique) disponibles. Cette façon de faire est très grossière car :

- d’une part, les aléas sont loin d’être tous gaussiens ;
- d’autre part, sommer les puissances maximales est peu réaliste, en particulier pour l’hydraulique. Nous pouvons en effet considérer des usines produisant à leur maximum alors qu’il n’y a aucune eau disponible en raison des contraintes d’influencement des vallées. La marge calculée est donc forcément un majorant de la marge réelle. Or, une marge trop importante pourrait conduire à des situations où nous estimerions qu’il est inutile de déclencher des décisions de couverture alors qu’en réalité la marge est insuffisante.

Pour construire cette marge, des scénarios de demande peuvent être obtenus à l’aide de modèles existants plus réalistes qu’une loi gaussienne. Il est donc important de concentrer les efforts sur l’amélioration de la capacité maximale de production.

La capacité maximale de production thermique : bon nombre de contraintes dynamiques n’étant pas considérées à l’horizon hebdomadaire, nous proposons de faire pour le thermique la même approximation que dans les pratiques actuelles : la capacité maximale de production thermique est la somme des productions maximales des centrales disponibles. Cependant, pour prendre en compte les pannes des centrales thermiques, nous avons introduit un modèle d’indisponibilité basé sur une loi exponentielle. En effet, si la centrale est opérationnelle aujourd’hui, le temps passé jusqu’à la prochaine panne peut être représenté par une loi exponentielle. Après une panne, une période de maintenance doit être respectée. Pour le moment, cette durée de réparation est fixe bien qu’elle pourrait être également aléatoire. Tous les paramètres ont été estimés en utilisant

les données historiques de pannes. Pour un tirage de pannes, nous calculons ainsi un scénario de capacité maximale de production thermique sur la période hebdomadaire.

La capacité maximale de production hydraulique : l'estimation de la capacité maximale de production hydraulique est un problème difficile pour deux raisons. D'une part, la capacité de production en un certain instant dépend de la gestion passée de la vallée (problème combinatoire). D'autre part, le nombre de vallées à prendre en compte (environ 40) amplifie cette nature combinatoire du problème. L'introduction de l'aléa d'hydraulicité complexifie cette estimation, notamment d'un point de vue numérique.

Pour calculer la capacité maximale hydraulique à l'instant t , nous proposons de résoudre le problème de gestion hydraulique (4.1) en incitant fortement la vallée à produire de l'électricité en introduisant un coût très élevé à ce pas de temps. Cette approche permet notamment de prendre en compte la satisfaction des niveaux de volume et le débit de la vallée, ce qui empêche la vallée de produire à outrance comme l'approche utilisée actuellement (somme des puissances maximales de toutes les usines hydro-électriques disponibles). Cependant, si nous supposons M hypothèses de prix (pour avoir M niveaux réalisables de capacités maximales), nous devons résoudre au premier pas de temps M problèmes d'optimisation. A partir de ces M capacités, nous devons faire de même à l'étape 2 ce qui conduit à M^2 problèmes d'optimisation. Si nous avons K pas de temps et V vallées, le nombre d'optimisations à effectuer est donc M^{KV} , ce qui peut être assez rapidement élevé. Pour simplifier cette combinatoire, nous résolvons le problème de gestion hydraulique sur 144 pas de temps mais en augmentant artificiellement les coûts marginaux aux heures de pointe de chaque jour. Nous obtenons ainsi une valeur moyenne de la capacité maximale de production hydraulique. Cette optimisation doit être répétée pour chaque tirage d'apports hydrauliques. La simplification du passage aux heures de pointe a pour principale conséquence de restreindre le problème de couverture uniquement aux heures de pointe. Mais il s'agit d'une hypothèse cohérente puisque, si la marge est suffisante à l'heure de pointe, elle le sera à toute heure de la journée. Nous introduisons aussi une sur-couverture par ce biais...

Les apports hydrauliques sont obtenues à l'aide d'une loi normale bi-tronquée dont les paramètres (bornes et écart-type) sont fonction de la saison : une saison pluvieuse aura plus de volatilité contrairement à une période de sécheresse.

Génération des scénarios de marge : pour générer un scénario de marge, il suffit de générer un scénario de demande, un scénario de capacité maximale de production hydraulique et un scénario de capacité maximale de production thermique et de les combiner entre eux. Nous pouvons obtenir ainsi autant de scénarios de marge que nous souhaitons.

Ce travail a été effectué en collaboration avec l'équipe du projet "Gestion Hebdomadaire de la production électrique". La procédure mise en place est une approche originale. Ce calcul a été intégré à un outil déjà utilisé en exploitation.

Formulation du problème en boucle ouverte

Le problème général de couverture contre le risque physique de défaillance peut s'écrire sous la forme générale suivante :

$$\begin{aligned} \min_{x \in X} \quad & \mathbb{E}[c(x, \xi)] \\ & \mathbb{P}[\tilde{M}_k(x, \xi) \geq M_S] \geq 1 - p, \quad \forall k = 1, \dots, K. \end{aligned} \quad (1.6)$$

Nous cherchons à minimiser l'espérance des coûts $c(x, \xi)$ de couverture, correspondant aux décisions x et à l'aléa ξ affectant le système. De plus, les décisions doivent satisfaire un ensemble de contraintes techniques X et assurer avec une probabilité supérieure à $1 - p$ un niveau de marge suffisant à chaque pas de temps k .

Dans la formulation (1.6), la marge dépend des décisions de couverture. En effet, un déplacement d'arrêt de centrale en maintenance modifie forcément la marge de production. Nous avons vu que le calcul de marge était loin d'être aisé, en particulier pour l'hydraulique. En effet, si nous intégrons des commandes ayant un impact sur le parc de production, nous devons intégrer le calcul de marge à l'intérieur du problème d'optimisation. Dans le cadre de l'hydraulique, cela revient à intégrer à l'intérieur de la contrainte en probabilité la résolution du problème de gestion de 40 vallées hydrauliques avec des apports incertains. Autant dire que ce problème est très difficile. Pour simplifier cela, nous ne considérerons dans un premier temps que les commandes financières.

Elles sont de deux types :

- les achats (ventes) sur les Marchés. Les produits concernés suivants sont disponible en Pointe et en Base :
 - les Day Ahead : décision d'achat (ou vente) en j pour livraison en $j + 1$;
 - les Week Ahead : décision d'achat (ou vente) en un jour ouvrable de la semaine S pour livraison égale sur tous les jours ouvrables de la semaine $S + 1$;
 - les Week-End Ahead : décision d'achat (ou vente) en un jour ouvrable de la semaine S pour livraison égale sur tout le week-end suivant.
- les options tarifaires : ce sont des contrats passés avec certains clients assurant un effacement de la production. Elles sont déclenchées en j pour effacement en $j + 1$.

Ces commandes n'influent pas directement la marge, celle-ci peut être calculée sous forme de scénarios à l'aide de la méthodologie présentée auparavant. Le problème à résoudre s'écrit donc :

$$\begin{aligned} \min_{x \in X} \quad & \mathbb{E}[c(x, \xi)] \\ & \mathbb{P}[\tilde{M}_j + x \geq M_S] \geq 1 - p, \quad \forall j = 1, \dots, J. \end{aligned} \quad (1.7)$$

Comme la marge n'est évaluée désormais qu'à l'heure de pointe, le problème (1.7) contient donc autant de contraintes que de jours d'étude. Si nous supposons que les variables du problème sont en boucle ouverte (elles ne sont dépendantes que du temps), donc anticipatives, et comme l'unique variable aléatoire au sein de la contrainte en probabilité (la marge) est scalaire et séparable (non en produit d'une variable), nous pouvons transformer chaque contrainte en probabilité par une contrainte linéaire par calcul du quantile à $1 - p$ de la distribution de probabilité de la marge [43]. Nous avons ainsi :

$$\mathbb{P}[\tilde{M}_j + x \geq M_S] \geq 1 - p, \quad \forall j = 1, \dots, J \Leftrightarrow q_j(1 - p) + x \geq M_S, \quad \forall j = 1, \dots, J,$$

1.3. SYNTHÈSE DES RÉSULTATS

où $q_j(1-p)$ est le quantile de la distribution de marge au jour j avec un niveau de $1-p$. Comme les décisions d'achats/ventes sont des actions binaires, le problème à résoudre est un problème d'optimisation linéaire avec variables mixtes. La taille du problème est tout à fait raisonnable en raison du faible nombre de contrats à prendre en compte et du faible nombre de pas de temps.

Nous avons testé cette approche sur une période de 12 jours. La situation est résumée sur la figure 5.2 : la marge initiale estimée n'est pas suffisante sur toute la seconde semaine avec même de la défaillance du mardi au jeudi (barres vertes). Nous considérons les produits suivants : les futures (day/week/week-end ahead) disponibles en achats et/ou en vente, en base et en pointe (sauf le week-end ahead uniquement en base) et 10 options tarifaires différentes (puissance d'effacement et coûts différents). La marge de sécurité a été fixée à 3000MW pour chaque jour et le niveau de probabilité p est fixé à 1%. Nous supposons que les prix de marché sont déterministes (prix moyens). Résoudre ce problème d'optimisation à variables mixtes (570 contraintes et 620 variables) prend moins d'une seconde sur un ordinateur personnel de 1.6 GHZ et en utilisant le kit libre de modélisation et de résolution GLPK. Les produits financiers permettent de satisfaire les niveaux de marge voulus en chaque jour comme le montre la figure 5.2. La figure 5.3 montre les commandes financières utilisées. Le coût de cette couverture est de 21 M€. Pour donner un ordre d'idée, un week ahead base de 3000MW à 50 €/MWh coûte 25.2 M€, ce qui illustre que la couverture obtenue est financièrement intéressante.

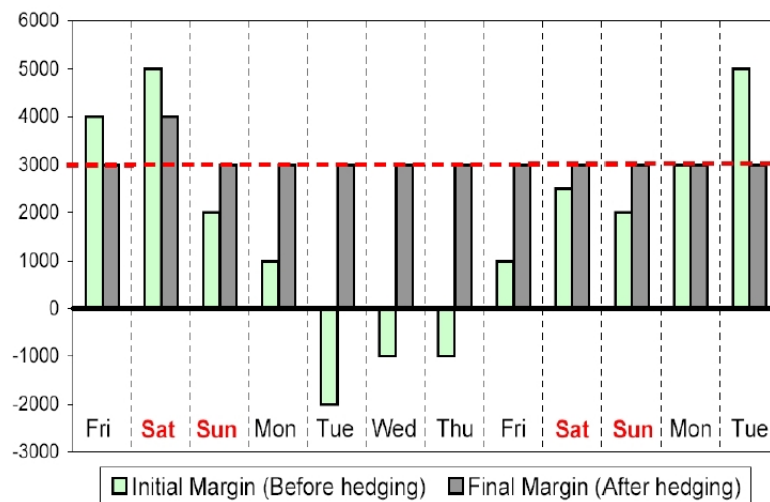


FIG. 1.4 – Marge de production avant et après la couverture

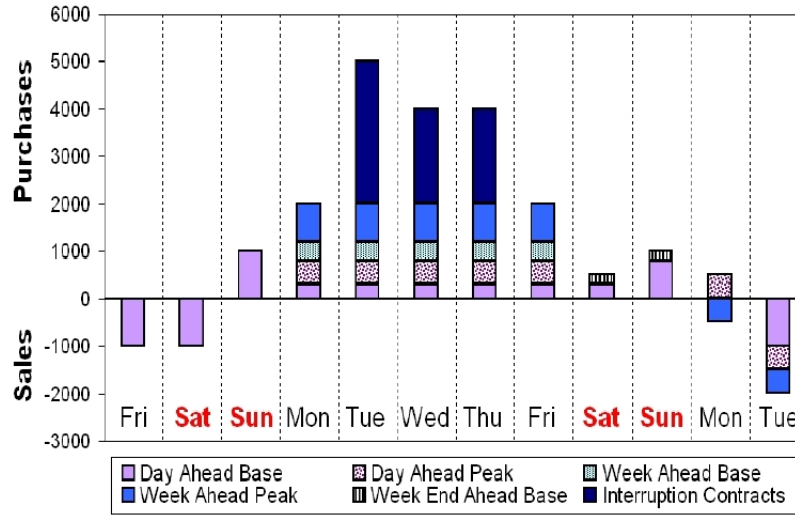


FIG. 1.5 – Commandes de couverture obtenues après optimisation

La procédure que nous proposons de mettre en place a été implémentée et donne lieu à une maquette (Matlab avec appel du kit de modélisation libre GLPK pour la résolution des problèmes linéaires mixtes) pouvant servir de démonstration aux utilisateurs.

Une approche pour la résolution du problème en boucle fermée : la Programmation Stochastique avec Règles de Décision Constantes par Morceaux

L'approche précédente est anticipative car elle suppose connu le futur dès le début de l'horizon. Une formulation en boucle fermée, avec des commandes dépendantes des réalisations de l'aléa, est certes plus réaliste mais elle introduit des difficultés supplémentaires. Dans notre cas, l'introduction des commandes en boucle fermée est assez problématique car effectuée dans des contraintes en probabilité. L'aléa n'est a priori plus séparable et la transformation en des contraintes linéaires n'est plus possible. De façon générale, il n'existe pas d'approches dans la littérature pour résoudre des problèmes multi-étapes dans des contraintes en probabilité. En effet, dans le cas statique classique, les contraintes en probabilité sont généralement non convexes voire même non connexes. Il existe aucun théorème pouvant nous en convaincre mais il est très probable que ces contraintes avec variables en boucle fermée conservent ces problèmes de non-convexité. Même en supposant que les contraintes en probabilité sont remplacées par des contraintes presque sûres, la difficulté demeure majeure.

Pour résoudre ce problème, nous avons choisi d'appliquer la programmation stochastique avec règles de décision constantes par morceaux, présentée en 6.1. Cependant, cette heuristique a été initialement conçue sur des problèmes sans contrainte sur l'état et avec un aléa unidimensionnel. Nous avons dû proposer des extensions pour appliquer l'heuristique à notre problème. En effet, plusieurs aléas interviennent : la marge et les prix de marché. Nous proposons donc une approche pour l'étendre au cas d'un aléa multidimensionnel. Pour cela, nous supposons que chaque dimension de l'aléa est indépendante. Nous procédons à la même étape de construction

d'ensembles de décision pour chacune des dimensions de l'aléa et nous y imposons de la même sorte des décisions uniques. Nous n'imposons donc plus des décisions v_t^k mais des décisions $v_t^{k,m}$, où chaque m est une dimension différente de l'aléa. Ainsi, quand nous considérons un scénario quelconque σ de dimension M , la règle de décision finale est la somme des décisions à prendre par rapport à chaque dimension de l'aléa :

$$u_t(\sigma_t) = \sum_{m=1}^M \sum_{k=1}^{|I_t^m|} \mathbb{1}_{\{\sigma_t \in I_t^{k,m}\}} v_t^{k,m}, \quad \forall t = 2, \dots, T-1. \quad (1.8)$$

Nous pouvons agir de même pour K experts différents et obtenir une solution finale qui prenne en compte leurs K avis de la même façon que nous l'avons présentée auparavant.

Nous avons d'autre part dû considérer que les contraintes en probabilité sont satisfaites sûrement ($p=0$). Nous supprimons ainsi la complexité de ces contraintes mais, en retour, nous n'avons aucune garantie sur la satisfaction des contraintes en simulation. Pour forcer cette satisfaction sur l'état en simulation, nous avons rajouté des coefficients de sécurité en optimisation. Mais ils ne doivent pas être trop élevés pour ne pas détériorer le coût moyen.

Pour construire les partitions, nous avons testé deux heuristiques différentes. Une première heuristique, "Variable Neighborhood Decomposition Search" (VNDS) introduite par Mladenovic et Hansen [40], a été utilisée. Les problèmes étaient parfaitement résolus mais dans des temps de calculs relativement importants (jusqu'à 9H de calculs!). Notre but n'étant pas forcément d'avoir une solution très précise mais plutôt une solution correcte en un minimum de temps, une nouvelle heuristique a été proposée dans [59]. Elle est basée sur la relaxation lagrangienne du problème de p -médianes, comme dans [31], et sur une technique d'arrondi aléatoire, comme dans [52]. Cet algorithme permet de reconstruire une solution primale du problème de p -médianes relâché (i.e. sans contrainte d'intégralité). Cette solution fractionnaire est interprétée comme une probabilité qu'un point (sous-scénario) soit choisi comme médiane (sous-scénario de référence). Sur cette base, nous construisons un algorithme aléatoire de tirages successifs de p médianes (p sous-scénarios de référence). Pour un tirage particulier, la partition autour des p sous-scénarios de référence se fait par affectation au plus proche sous-scénario de référence. L'opération est répétée un grand nombre de fois et la meilleure solution est retenue. La solution de cette approche est certes moins précise (de l'ordre de 5%) mais permet de réduire les temps de calculs de manière considérable (au moins de moitié en général). Nous pouvons noter que ces calculs pourraient être massivement parallélisés puisque toutes ces constructions de partitions sont indépendantes.

Pour tester cette heuristique, nous avons choisi de travailler sur une version simplifiée de (1.7), de petite taille (10 jours, un contrat week-ahead, un contrat day-ahead et un stock d'options tarifaires) et avec uniquement des variables continues (au lieu de mixtes). Les résultats sont assez encourageants puisque, pour certains paramétrages (nombre et taille d'experts suffisants pour que les contraintes soient satisfaites sur un nombre de scénarios suffisant en optimisation), le risque 1% est vérifié et nous dégageons des gains substantiels. Les gains peuvent même être meilleurs que l'optimisation robuste ou la formulation en boucle ouverte (jusqu'à 20% d'amélioration). L'introduction de coefficients de sécurité aide effectivement à la satisfaction des contraintes

de marge, moyennant une réduction des gains (pouvant être élevée, de l'ordre de 20%).

Pistes d'améliorations

Cette partie étant très exploratoire, elle ouvre de nombreuses pistes pour la suite de cette thèse. Nous pouvons les développer par thèmes :

Calcul de la marge : le calcul de la marge proposé est une véritable évolution par rapport aux approches existantes à l'heure actuelle. Cependant, il peut encore être amélioré :

- cas de la capacité maximale thermique : nous calculons la capacité maximale comme la somme des puissances maximales des centrales disponibles. Or, en raison des paliers de production, des durées minimales de palier, les courbes de démarrage, toutes les centrales ne peuvent produire à leur maximum. Nous pouvons donc imaginer résoudre le problème thermique par programmation dynamique pour prendre en compte ces particularités ;
- cas de la capacité maximale hydraulique : la capacité maximale à un pas de temps t dépend de l'utilisation qui a été faite de l'eau sur les pas de temps précédents. Cette information étant résumée dans le niveau de volume t , ce processus est Markovien. L'utilisation de cet état de fait peut permettre la réduction de la combinatoire du problème. D'autre part, pour calculer la marge hydraulique à t_0 sur les J prochains jours, il est possible d'utiliser l'optimisation robuste. Si nous fixons une chronique de prix déterministe, il suffit de résoudre le problème robuste à l'aléa d'hydraulicité en boucle fermée. Or, nous avons montré dans la première partie de thèse comment résoudre un tel problème et les performances prometteuses de l'approche.

Formulation en Boucle ouverte : jusqu'à présent, nous évaluons des scénarios de marge à état de parc initial fixé à partir desquels nous résolvons le problème d'optimisation de la couverture avec uniquement des actifs financiers. Or, nous pouvons imaginer pouvoir utiliser le parc de production comme leviers de couverture. Nous pouvons notamment déplacer des arrêts en maintenance de centrales ou utiliser le parc hydraulique. Cependant, l'utilisation de ces commandes modifie automatiquement la valeur de la marge. Or, puisque les scénarios de marge sont calculés préalablement à l'optimisation, il faut revoir la méthodologie et même intégrer le calcul de marge à l'intérieur du problème. Une expérimentation de cette thèse a montré qu'en intégrant le calcul de marge thermique (avec une indisponibilité fixée) dans le problème d'optimisation de l'exemple des figures 5.2 et 5.3, nous réduisons le coût de couverture de 14% en déplaçant une centrale, 19% en déplaçant deux centrales. L'enjeu économique est donc important.

Cependant cette intégration de la marge dans le problème de l'optimisation est assez difficile à mettre en œuvre :

- cas de la capacité maximale thermique : le calcul de la capacité maximale thermique telle qu'elle est calculée à l'heure actuelle peut se faire aisément si l'indisponibilité est fixée. Il s'agit juste de rajouter des variables supplémentaires. Mais comme l'aléa (marge hydraulique) est toujours séparable et scalaire, nous pouvons toujours transformer les contraintes en probabilité en contraintes linéaires. Par contre, si nous souhaitons rajouter un modèle d'indisponibilité plus complexe, l'aléa n'est plus séparable et est multiplicatif avec les variables de décision. D'ailleurs rien n'assure que désormais l'ensemble de faisabilité soit convexe dans ce cas ;
- cas de la capacité maximale hydraulique : pour intégrer le calcul de la capacité maximale hydraulique, il faut intégrer un problème d'optimisation linéaire à l'intérieur de la contrainte

en probabilité. C'est donc un problème extrêmement difficile.

Formulation en Boucle fermée : cette formulation ouvre de très nombreuses perspectives que cette thèse n'a pas pu couvrir par manque de temps. Nous pouvons citer les pistes suivantes qui sont dans le prolongement direct des travaux de thèse et qui sont intéressants :

- continuer à exploiter la SPSDR. Les résultats obtenus sont des résultats préliminaires et il serait très intéressant de continuer à développer cette approche ;
- mettre en place de nouvelles méthodes pour comparer les résultats. Nous pouvons notamment appliquer l'optimisation robuste avec ajustements linéaires (comme pour l'hydraulique). Cependant, pour cela, il est nécessaire d'avoir à disposition un modèle auto-régressif de la marge. Sinon les règles de décision complexifient le problème. Une autre façon d'appliquer l'optimisation robuste est de la combiner avec la programmation stochastique. Les commandes deviennent des règles de décision fonction de l'aléa de prix. La contrainte de marge est robustifiée par rapport aux aléas de marge et de prix. Par contre, la fonction économique est gérée par programmation stochastique, ce qui permet de résoudre le problème d'aléas multiplicatifs. Cette approche semble encourageante et mériterait d'être expérimentée.

1.4 Pour résumer : Apports principaux de la thèse

1.4.1 Aspects méthodologiques de la thèse

Dans cette thèse, nous avons cherché à étudier et appliquer deux approches intéressantes et originales dans le secteur de la gestion de production énergétique : l'optimisation robuste et la programmation stochastique avec règles de décision constantes par morceaux.

L'optimisation robuste, nous le rappelons, permet de s'assurer de la satisfaction des contraintes pour toute réalisation de l'aléa au sein d'un ensemble d'incertitude choisi par le décideur. Dans la première partie de cette thèse, après avoir analysé ces approches robustes et donc défini le cadre théorique général, nous nous sommes concentrés à les appliquer dans un contexte industriel. La problématique choisie est la gestion d'une vallée hydraulique. Un des aléas essentiels affectant la gestion hydraulique est la grande variabilité des apports hydrauliques aux réservoirs des vallées. Nous pouvons faire des prévisions de ces apports hydrauliques sur les deux semaines à venir et estimer les écarts maximaux sur la tendance associée. Les décisions optimales construites avec une approche déterministe ne sont pas robustes à cet aléa et, ainsi, conduisent à des violations de contraintes. Dans ce contexte, nous proposons d'aborder ce problème à l'aide de méthodes d'optimisation robuste en vue d'obtenir des stratégies robustes de production. Ce choix, déjà justifié par la nécessité de mettre en oeuvre une approche nouvelle, de la maîtriser et de l'évaluer par rapport à des techniques existantes, se justifie aussi par les enjeux importants associés à la définition d'une gestion hydraulique robuste aux aléas à l'horizon hebdomadaire. Une telle approche robuste a été testée sur une vallée suffisamment représentative d'une partie des vallées hydrauliques françaises. Les approches robustes ont été comparées avec une approche dite Déterministe à Révision Périodique (DRP), proche des pratiques actuelles en exploitation. Les principaux résultats montrent que les modèles robustes permettent de diminuer notablement les violations de contrainte de volume (de l'ordre de 75 à 95% dans notre exemple) jusqu'à les rendre nulles au prix d'une dégradation de la fonction économique de seulement 0,5% (par

rapport à la borne maximale de gain, issue de l'optimisation en information parfaite) et d'un accroissement raisonnable de la taille du problème. Les résultats obtenus par l'approche robuste sont donc satisfaisants et permettent d'envisager à terme une prise en compte opérationnelle des aléas dans le problème de gestion de la production hydraulique.

La deuxième approche étudiée est la programmation stochastique avec règles de décision constantes par morceaux (SPSDR). Nous rappelons que la SPSDR permet de palier au problème de croissance exponentielle de l'arbre en programmation stochastique classique en supprimant la contrainte de raffinement. De plus, l'utilisation d'experts indépendants permet d'avoir une bonne estimation de la solution tout en réduisant les temps de calculs. Cette approche est très originale mais restreinte dans sa création aux aléas unidimensionnels et à des états sans contrainte. Nous proposons dans cette thèse deux extensions à l'heuristique initiale de [59] : le cas d'un aléa multidimensionnel et l'ajout de bornes sur les états. La gestion du cas multidimensionnel consiste à faire une partition des scénarios en ensembles de décision pour chacune des dimensions de l'aléa. Une optimisation indépendante par rapport à chaque dimension permet d'avoir des règles, qui constituent différents points de vue au même problème selon l'aléa pris en compte. La règle de décision finale (pour un expert) est finalement la combinaison des règles obtenues indépendamment sur chaque dimension de l'aléa. D'autre part, nous pouvons forcer la satisfaction des contraintes sur les états en introduisant des coefficients de sécurité. Les premiers résultats de cette approche sur des problèmes classiques sont relativement encourageants [59]. Cette approche a été appliquée à la résolution du problème multi-étapes de couverture contre le risque de défaillance physique sur lequel nous reviendrons dans la sous-section suivante. Nous montrons notamment que, pour certains paramétrages (c-à-d avec des experts suffisamment nombreux et des experts suffisamment grands pour que les contraintes soient vérifiées sur le plus de scénarios possibles), les stratégies obtenues permettent en simulation de satisfaire le critère de risque à 1%. De plus, les performances économiques sont meilleures qu'avec une approche robuste ou une approche en boucle ouverte (jusqu'à 20% de gains). Ces premiers résultats montrent le réel intérêt de cette approche.

Nous avons ainsi testé deux approches originales à deux problématiques différentes. Toutes deux fournissent des résultats très convaincants, constituant ainsi un véritable apport de cette thèse.

1.4.2 Modélisation de problèmes industriels

Bien que nous ayons cherché à étudier l'intérêt de techniques originales, la volonté de cette thèse est avant tout d'apporter des solutions à des problèmes industriels concrets. La programmation robuste nous apporte déjà des premières réponses à la gestion hydraulique sous aléa. Cependant, pour une meilleure gestion du risque de défaillance physique, nous nous sommes replacés dans le cadre général de prise de décisions (financière et physique) de couverture à l'horizon hebdomadaire. Il s'agit donc de modéliser un véritable problème industriel à forts enjeux économiques.

Nous n'allons pas re-présenter cette problématique introduite dans la première section de ce chapitre introductif. Partant d'aucun existant, nous avons dû dans un premier temps analyser le problème afin de proposer une formulation. Il nous a semblé naturel d'étudier le problème

global, optimisant conjointement la production et les risques. Idéalement, les marchés devraient être intégrés à l’optimisation du planning de production, tout comme ce critère de risque. En effet, le producteur optimiserait ainsi conjointement son parc de production et ses achats sur les marchés, ce qui lui permettrait de satisfaire son critère de risque à moindre coût, notamment à l’aide de moyens souvent moins chers que les prix de marché (comme l’hydraulique). Cependant, résoudre un problème d’optimisation avec un parc de production modélisé finement (introduisant notamment de la non-convexité à cause de paliers de production, de coûts de démarrage,... car nous sommes au court-terme), de taille importante et avec un aléa multidimensionnel est très difficile.

Après avoir écrit ce problème général, nous proposons dans cette thèse une alternative sous-optimale permettant toutefois d’obtenir une politique de couverture satisfaisant ce risque de défaillance physique. Dans un premier temps, le problème de planification est résolu en déterministe à l’aide de prévisions moyennes pour chacun des aléas. Dans un second temps, à partir de cet état de parc “optimal”⁵ ainsi obtenu, nous optimisons les décisions de sécurité en vue d’assurer une marge de production satisfaisante dans les $100 - p\%$ cas, la marge étant définie comme la différence entre l’offre disponible maximale et la demande. L’offre maximale étant soumise aux aléas du parc de production (indisponibilité et hydraulité) et comme la demande est par nature aléatoire, la marge est elle-même un processus stochastique.

Cette deuxième étape peut en fait se formuler sous la forme d’un problème d’optimisation stochastique avec contrainte en probabilité. Dans le cas dit de boucle ouverte, c’est à dire anticipatif, où toutes les commandes de l’horizon sont décidées une fois pour toute à l’instant initial, les contraintes peuvent être transformées en contraintes linéaires par calcul de quantile, l’aléa de marge étant séparable. Mais pour cela, il est nécessaire de connaître la distribution de la marge et savoir modéliser cet aléa (par exemple sous forme de scénarios). Nous proposons dans cette thèse une technique de calcul de la marge. Nous nous basons notamment sur l’hypothèse que seule l’heure de pointe de chaque journée doit être prise en compte : si le producteur dispose suffisamment d’énergie pour satisfaire la pointe de production alors il en disposera de suffisamment pour satisfaire la demande aux autres heures. Le calcul de l’offre maximale thermique est simplifié : nous la supposons égale à la somme des productions maximales de toutes les centrales disponibles. Nous ne prenons donc pas en compte le caractère discret (paliers de production) des décisions thermiques. Pour l’offre maximale hydraulique, nous résolvons le problème de gestion de chaque vallée en imposant une chronique de prix pour laquelle les heures de pointe sont fortement rémunérées. Nous créons ainsi des scénarios de marge qui sont fournis au modèle d’optimisation et qui permettent de calculer les quantiles. Au final, le problème de couverture à résoudre est un problème linéaire de taille tout à fait raisonnable.

Dans le cas de la boucle fermée, le problème est bien plus complexe. La littérature ne propose pas d’approches permettant la résolution de problèmes avec des variables en boucle fermée dans des contraintes en probabilité. Faute d’approche satisfaisante, nous avons testé la programmation stochastique avec règles de décision constantes par morceaux, approche qui n’a initialement pas été conçue pour un tel problème. Nous montrons notamment que, pour certains paramétrages (c’est à dire avec des experts suffisamment nombreux et des experts suffisamment grands pour

⁵gestion du parc optimale pour les prévisions moyennes de l’aléa à l’horizon hebdomadaire, mais gestion peu robuste à la moindre variation de n’importe quel aléa soumis au système.

que les contraintes soient vérifiées sur le plus de scénarios possibles), les stratégies obtenues permettent en simulation de satisfaire le critère de risque à 1%. De plus, les performances économiques sont meilleures qu’avec une approche robuste ou une approche en boucle ouverte (jusqu’à 20% de gains).

La modélisation proposée est une première alternative permettant d’optimiser la couverture contre le risque de défaillance physique. Bien que sous-optimale mathématiquement, elle apporte une plus-value certaine pour le contrôle de la couverture. D’autant plus que cette approche est relativement similaire à certains processus “métier” appliqués chez certains producteurs. Elle a donc le mérite d’optimiser et d’automatiser ces processus. Nous apportons ainsi une analyse et une formulation du problème, une justification de pratiques appliquées en exploitation, une modélisation pour les aléas comme la marge et des techniques de résolution en boucle ouverte et en boucle fermée. Enfin, cette partie étant exploratoire, nous avons également ouvert de nombreuses pistes de réflexion et d’amélioration à la suite de cette thèse.

1.4.3 La validation par simulation

Outre l’originalité des approches utilisées et la modélisation du problème industriel de couverture, cette thèse met également en avant un aspect supplémentaire : celui de la validation par simulation.

Les approches étudiées durant nos études ne sont pas forcément fondées sur des résultats théoriques assurant la convergence des approches. En optimisation robuste, il existe des théorèmes bornant la probabilité de non-satisfaction de la contrainte quand les variables sont en boucle ouverte. Dans le cas de la boucle fermée ou même de la programmation stochastique avec règles de décision constantes par morceaux, aucun théorème ne permet de justifier que l’approche utilisée est efficace. Pour cela, il est donc nécessaire d’évaluer la solution obtenue par de la simulation numérique.

Le principe est simple : lors de la phase d’optimisation (et donc d’obtention des stratégies), nous avons utilisé une certaine connaissance de l’aléa (moyenne et écart type pour l’aléa d’hydraulicité dans le cas de l’optimisation robuste, scénarios de marge et de prix dans le cas de la programmation stochastique avec règles de décision constantes par morceaux). Dans la phase de simulation, nous générons un très grand nombre de scénarios réalistes et nous y appliquons la stratégie obtenue. Dans le cadre de règles de décision affines en l’aléa, nous pouvons avoir notamment à reconstruire les commandes. Nous pouvons également, comme pour la gestion hydraulique, simuler un processus métier rendant la simulation plus “intelligente”, corrigeant même les décisions à l’aide d’heuristiques complémentaires. Par ce procédé, nous pouvons calculer divers indicateurs tels que le coût moyen, l’écart type des coûts, la distribution des coûts, la satisfaction des contraintes,... Cette approche est très expérimentale mais a le mérite de valider la pertinence de la solution proposée, et donc de l’approche appliquée. Elle permet aussi de s’étalonner par rapport à des processus métier existants qui sont généralement la référence de comparaison des décideurs. Par contre, les temps de calculs pour la simulation peuvent être non négligeables (notamment si des règles de décision doivent être construites) et peuvent contraindre le nombre de scénarios utilisés à être, peut être, insuffisant pour valider la méthode de façon rigoureuse. Mais faute de résultats rigoureux, cette validation par simulation paraît être nécessaire pour

convaincre un décideur d'appliquer une stratégie donnée.

Nous pouvons rajouter enfin que, même pour des approches classiques fondées sur des théorèmes de convergence (gradient stochastique, programmation stochastique), l'application de la simulation n'est pas dénuée de sens. Pourtant, dans la majorité des articles de la littérature consacrée à ces approches, l'optimisation est effectuée et les résultats en découlant sont mis en avant directement sans aucune validation. Par exemple, le coût optimal obtenu en espérance sur un arbre d'aléa. Mais ces résultats sont entièrement dépendants de la modélisation de l'aléa dans le modèle. Pour reprendre l'exemple de l'arbre de scénarios, le coût optimal ne l'est que pour les scénarios utilisés en optimisation, surtout si l'arbre est réduit pour conserver des temps de calculs raisonnables. Or, il est irréaliste de croire que les scénarios utilisés pour construire l'arbre seront suffisamment représentatifs et nombreux pour refléter l'ensemble des états de la nature possibles. Par conséquent, qu'en est-il du coût pour de nouveaux scénarios, différents de ceux de l'optimisation ? La stratégie obtenue est elle vraiment performante ? Les contraintes techniques "dures" demeurent-elles toujours satisfaites ? Ce sont le type de questions que nous pouvons en toute légitimité nous poser et dont la phase de simulation peut apporter des éléments de réponse.

1.5 Publications et communications orales

1.5.1 Communications orales

Cette thèse a donné lieu à de nombreuses communications orales. L'application de l'optimisation robuste à la gestion hydraulique a été exposée à la conférence internationale "Mathematical Programming" [8], au 3ème Cycle Romand de Recherche Opérationnelle [10, 9] et à la conférence jointe ROADEF-FRANCORO [6]. Elle a également été présentée dans le cadre de séminaire à l'Institut Français du Pétrole [7].

La modélisation du problème de couverture contre le risque de défaillance physique a été présentée à la conférence "Power Systems Management Conference" [3] et à la conférence "International Conference on Stochastic Programming" [2].

Enfin, l'application de la programmation stochastique avec règles de décision constantes par morceaux au problème de couverture contre le risque de défaillance physique en boucle fermée a également été présentée à la conférence "International Conference on Stochastic Programming" [4].

1.5.2 Articles et chapitres de livre

Dans le cadre du "Handbook on Uncertainty and Environmental Decision Making", deux chapitres ont été proposés : un chapitre d'introduction à l'optimisation robuste linéaire [62] et un chapitre sur son application à la gestion d'une vallée hydraulique [11].

D'autre part, la modélisation du problème de couverture contre le risque de défaillance physique et le calcul de la marge ont donné lieu à un article accepté pour publication dans IEEE Power

Systems [64].

1.5.3 Notes internes

Le chapitre 3 d'introduction à l'optimisation robuste linéaire a donné lieu à une note d'initiation, interne à EDF [61]. Son application à la gestion hydraulique a donné lieu à la note [5].

Tout le processus de calcul de la marge (première partie du chapitre 7) a été présenté dans la note [63].

La formulation en boucle ouverte du problème de couverture contre le risque de défaillance sera diffusée dans la note [12].

Bibliographie

- [1] E. Adida and G. Perakis. A Robust Optimization Approach to Dynamic Pricing and Inventory Control with no Backorders. *Mathematical Programming*, 107(1-2) :97 – 129, 2006.
- [2] L. Andrieu, R. Apparigliato, A. Lisser, and A. Tang. Stochastic Optimization under Risk Constraints : Application to Hedging Problem in Electrical Industry. *11th International Conference on Stochastic Programming, Vienna*, August 30th 2007.
- [3] L. Andrieu, R. Apparigliato, A. Lisser, and A. Tang. Stochastic Optimization under Risk Constraints : Application to Hedging Problem in Electrical Industry. *Power Systems Management Conference, Athens*, June 7th 2007.
- [4] R. Apparigliato, J. Thénier, and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : Application on a Hedging Problem in Electrical Industry. *11th International Conference on Stochastic Programming, Vienna*, August 30th 2007.
- [5] R. Apparigliato, J.-P. Vial, and R. Zorgati. Optimisation Robuste Linéaire - Application : Gestion Court Terme d'une Vallée Hydraulique. Technical report, EDF R&D, Dept OSIRIS, H-R32-2007-00658-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [6] R. Apparigliato, J.-P. Vial, and R. Zorgati. Gestion Hebdomadaire d'une Vallée Hydraulique par Optimisation Robuste. *Conférence conjointe FRANCORO-ROADEF, Grenoble (France)*, 23 Février 2007.
- [7] R. Apparigliato, J.-P. Vial, and R. Zorgati. Gestion Hebdomadaire d'une Vallée Hydraulique par Optimisation Robuste. *Séminaire scientifique de l'Institut Français du Pétrole, Rueil-Malmaison (France)*, 26 octobre 2007.
- [8] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. *19th International Symposium on Mathematical Programming, Rio de Janeiro (Brazil)*, August 2nd 2006.
- [9] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. *Third Romand Cycle of Operational Research, Zinal (Switzerland)*, March 5th 2007.

- [10] R. Apparigliato, J.-P. Vial, and R. Zorgati. An Original Approach for the Weekly Management of a Hydraulic Valley. *Third Romand Cycle of Operational Research, Zinal (Switzerland)*, March 7th 2006.
- [11] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. In F. S. Hillier, editor, *Handbook on Uncertainty and Environmental Decision Making*, International Series in Operations Research and Management Science. Springer, to appear (submitted).
- [12] R. Apparigliato and R. Zorgati. Couverture et Arbitrage sur les Marchés en Gestion Hebdomadaire de la Production Electrique. Partie II : Formulation du Problème de Couverture. Technical report, EDF R&D, Dept OSIRIS, 2008 (A paraître).
- [13] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending the Scope of Robust Optimization : Comprehensive Robust Counterparts of Uncertain Problems. *Mathematical Programming Series B*, 107(1-2) :63–89, 2006.
- [14] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-Ph. Vial. Retailer-Supplier Flexible Commitments Contracts : a Robust Optimization Approach. *Manufacturing and Service Operations Management*, 7 :248–271, 2005.
- [15] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming*, 99(2) :351–376, 2004.
- [16] A. Ben-Tal and A. Nemirovski. Robust Convex Optimization. *Mathematics of Operations Research*, 23 :769 – 805, 1998.
- [17] A. Ben-Tal and A. Nemirovski. Robust Solutions of Linear Programming Problems Contaminated with Uncertain Data. *Mathematical Programming Series A*, 88 :411 – 424, 2000.
- [18] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization : Analysis, Algorithms and Engineering Applications*. MPS-SIAM series on Optimization. Society for Industrial and Applied Mathematics and Mathematical Programming Society, 2001.
- [19] A. Ben-Tal and A. Nemirovski. Selected Topics in Robust Convex Optimization. *Mathematical Programming : Series A and B*, 112 :125–158, 2007.
- [20] D. Bertsimas, D. Pachamanova, and M. Sim. Robust Linear Optimization under General Norms. *Operations Research Letters*, 32 :510–516, 2004.
- [21] D. Bertsimas and M. Sim. Price of Robustness. *Operations Research*, 52 :35–53., 2004.
- [22] D. Bertsimas and A. Thiele. A Robust Optimization Approach to Inventory Theory. *Operations Research*, 54(1) :150–168, 2006.
- [23] Daniel Bienstock. Experiments in Robust Portfolio Optimization. Technical report, Center for Financial Engineering, Columbia University, January 2007.
- [24] G. C. Calafiore. Ambiguous Risk Measures and Optimal Robust Portfolios. Technical report (to appear in *SIAM Journal on Optimization* 2007), Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy, 2007.
- [25] G. C. Calafiore and L. El-Gahoui. On Distributionally Robust Chance-Constrained Linear Programs. *Journal of Optimization Theory and Applications*, 130 :1–22, 2006.
- [26] S. Ceria and R. Stubbs. Incorporating Estimation Errors into Portfolio Selection : Robust Portfolio Construction. *Journal of Asset Management*, 7(2) :109 – 127, 2006.

- [27] A. Charnes and W.W. Cooper. Chance Constrained Programming. *Management Science*, 6 :73–89, 1959.
- [28] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost Horizons and Certainty Equivalents : an Approach to Stochastic Programming of Heating Oil. *Management Science*, 4 :235–263, 1958.
- [29] X. Chen, M. Sim, and P. Sun. A Robust Optimization Perspective of Stochastic Programming. Working paper, NUS Business school, 2005.
- [30] D. Dentcheva and W. Römisch. Optimal Power Generation under Uncertainty via Stochastic Programming. *Stochastic Programming Methods and Technical Applications*, 458 :22–56, 1998.
- [31] O. du Merle and J.-P. Vial. Proximal ACCPM, a Cutting Plane Method for Column Generation and Lagrangian Relaxation : Application to the p-median Problem. Technical report, Logilab - HEC - University of Geneva, 2002.
- [32] J. Dupačová, N. Growe-Kuska, and W. Römisch. Scenario Reduction in Stochastic Programming. An Approach using Probability Metrics. *Mathematical Programming*, 95(3) :493–511, 2003.
- [33] L. El-Ghaoui and H. Lebret. Robust Solutions to Least-square Problems to Uncertain Data Matrices. *SIAM Journal of Matrix Analysis and Applications*, 18 :1035–1064, 1997.
- [34] S.-E. Fleten and T.K. Kristoffersen. Short-Term Hydropower Production Planning by Stochastic Programming , 2008.
- [35] K. Frauendorfer and J. Güssow. *Stochastic Multistage Programming in the Optimization and Management of a Power System*. Stochastic Optimization Techniques - Numerical Methods and Technical Applications. Springer-Verlag, 2002.
- [36] D. Goldfarb and G. Iyengar. Robust Portfolio Selection Problems. *Mathematics of Operations Research*, 28 :1 – 38, 2003.
- [37] N. Gröwe-Kuska, K.C. Kiwiel, M.P. Nowak, W. Römisch, and I. Wegner. Power Management in a Hydro-thermal System under Uncertainty by Lagrangian Relaxation. In A. Borghetti, C.A. Nucci, and M. Paolone, editors, *IMA Volumes in Mathematics and its Applications*, volume 128, pages 39–70. Springer-Verlag, New-York, 2002.
- [38] N. Gröwe-Kuska and W. Römisch. *Stochastic Unit Commitment in Hydro-thermal Power Production Planning*, chapter 30. MPS-SIAM Series in Optimization, 2005.
- [39] J. Güssow. *Power Systems Operation and Trading in competitive Energy Markets*. PhD thesis, University of St. Gallen, 2001.
- [40] P. Hansen and N. Mladenovic. Variable Neighborhood Search. *Computers and Operations Research*, 24 :1097–1100, 1997.
- [41] H. Heitsch and W. Römisch. Scenario Reduction Algorithms in Stochastic Programming. *Computational optimization and applications*, 24 :187–206, 2003.
- [42] H. Heitsch and W. Römisch. Generation of Multivariate Scenario Trees to Model Stochasticity in Power Management. *Computational Optimization and Applications*, 24(2-3) :187 – 206, February-March 2003.
- [43] R. Henrion. Introduction to Chance-Constrained Programming. *Tutorial paper for the Stochastic Programming Community Home Page*, 2004.

- [44] G. Iyengar. Robust Dynamic Programming. *Mathematics of Operations Research*, 30 :257 – 280, 2005.
- [45] O. Kariv and L Hakimi. An Algorithmic Approach to Network Location Problems. II : the p-medians. *SIAM Journal of Applied Mathematics*, 37(3) :539–560, 1979.
- [46] B. Mo, A. Gjelsvik, and A. Grundt. Integrated Risk Management of Hydro Power Scheduling and Contract Management. *IEEE Transactions on Power Systems*, 16(2) :216–221, 2001.
- [47] Y. Nesterov and J.-Ph. Vial. Confidence Level Solutions for Stochastic Programming. *Automatica*, 2008 (à paraître).
- [48] R. Nürnberg and W. Römissh. A Two-stage Planning Model for Power Scheduling in a Hydro-thermal System under Uncertainty. *Optimization and Engineering*, 3 :355–378, 2002.
- [49] A. Ouorou and J.-P. Vial. A Model for Robust Capacity Planning for Telecommunications Networks under Demand Uncertainty. In *submitted to the 6th International Workshop on Design and Reliable Communication Networks, DRCN*, 2007.
- [50] M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage Stochastic Optimization applied to Energy Planning. *Mathematical Programming*, 52 :359–375, 1991.
- [51] A. Prékopa and T. Szántai. On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake. *European Journal of Operations Research*, 3 :175–189, 1979.
- [52] P. Raghavan and C.D. Thompson. Randomized Rounding : a Technique for Provably Good Algorithms and Algorithmic Proofs. *Combinatorica*, 7(4) :365–374, 1987.
- [53] T.A. Rotting and A. Gjelsvik. Stochastic Dual Dynamic Programming for Seasonal Scheduling in the Norwegian Power System. *IEEE Transactions on Power Systems*, 7(1) :273–279, 1992.
- [54] A. Shapiro. Complexity of Two and Multi-stage Stochastic Programming Problems. Conference transparencies, 2005.
- [55] A. Shapiro. On Complexity of Multistage Stochastic Programs. *Operations Research Letters*, 34 :1 – 8, 2006.
- [56] A. Shapiro. Stochastic Programming Approach to Optimization under Uncertainty. Working paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332- Atlanta, Georgia 30332-0205, USA, 2006 (to appear in *Mathematical Programming*).
- [57] A. Shapiro and A. Nemirovski. On Complexity of Stochastic Programming Problems. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA, 2004.
- [58] A. L. Soyster. Convex Programming with Set-inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21 :1154–1157, 1973.
- [59] J. Thénié and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : a Heuristic Approach. *Automatica*, 2008 (to appear).

BIBLIOGRAPHIE

- [60] G. Unger. *Hedging Strategy and Electricity Contract Engineering*. PhD thesis, ETH, Zürich, 2002.
- [61] J.-P. Vial and R. Apparigliato. Optimisation Robuste Linéaire - Une Introduction Succinte. Technical report, EDF R&D, Dept OSIRIS, H-R32-2006-04801-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [62] J.-P. Vial, R. Apparigliato, and F. Babonneau. Introduction to Linear Robust Optimization. In F. S. Hillier, editor, *Handbook on Uncertainty and Environmental Decision Making*, International Series in Operations Research and Management Science. Springer, to appear (submitted).
- [63] R. Zorgati, R. Apparigliato, and J.-M. Luel. Couverture et Arbitrage sur les Marchés en Gestion Hebdomadaire de la Production Electrique. Partie I : Marge de Production et Risque Physique à 1%. Technical report, EDF R&D, Dept OSIRIS, H-R37-2007-01772-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [64] R. Zorgati, W. Van Ackooij, and R. Apparigliato. Supply Shortage Hedging : Estimating the Electrical Power Margin for Optimizing Financial and Physical Assets with Chance-Constrained Programming. *submitted to IEEE Transaction on Power Systems*, March 26th, 2008.

Chapitre 2

L'optimisation dans l'incertain

Sommaire

2.1	Les modèles statiques	52
2.1.1	Le gradient stochastique	53
2.1.2	L'approche par scénarios	54
2.1.3	L'aléa dans les contraintes	54
2.2	Les problèmes à deux étapes	56
2.3	Les problèmes à plusieurs étapes	60
2.3.1	La programmation stochastique sur arbre	60
2.3.2	La programmation dynamique et ses extensions	62
2.3.3	Les autres approches	64
2.4	Conclusions	64
	Bibliographie	65

En gestion de la production électrique (et c'est un exemple parmi tant d'autres), nous cherchons à trouver des solutions optimales au sens de critères économiques, sous un grand nombre de contraintes techniques et à divers horizons temporels. Mais comment être capable de prendre des décisions pour dans un an, voire même pour dans 15 jours, alors que de très nombreux événements imprévisibles peuvent avoir lieu : des anomalies climatiques, des pannes de centrales, des pics de consommation électrique non prévus, ... ? Evidemment, une décision prise à partir de mauvaises prévisions générera un planning qui ne correspondra pas à la situation réelle. Dans ce cas, soit nous sommes capables de corriger le tir, via des décisions de secours à des prix élevés, et nous désoptimisons complètement le problème ; soit nous n'en sommes pas capables et nous nous exposons à des situations extrêmes comme un black-out. Dans un contexte où les risques doivent être de mieux en mieux maîtrisés, il est donc nécessaire d'intégrer cette notion d'incertitude dans la modélisation. Pour cela, la nature des problèmes peut varier selon de nombreuses hypothèses : avons-nous une représentation probabiliste de l'aléa ? Le problème est-il statique ou dynamique ? Est-ce qu'une prise en compte moyenne des aléas est suffisante ? Voulons-nous des commandes toujours réalisables ou uniquement en moyenne ou en probabilité ? Voici quelques questions que nous devons nous poser avant de résoudre le problème. Dans ce chapitre, nous proposons une synthèse sur l'optimisation dans l'incertain afin de présenter tous les concepts appliqués et de présenter les principales approches utilisées.

2.1 Les modèles statiques

Introduisons quelques notations :

- $\chi \subset \mathbb{R}^n$ l'espace des variables de décision ;
- $X \subset \chi$ l'ensemble des décisions réalisables ;
- $x \in X$ un vecteur de décisions ;
- ω un élément de l'espace probabilisé Ω munie d'une probabilité P et d'une σ -algèbre \mathcal{F} ;
- une fonction objectif $F(x, \omega)$ mesurable et convexe en x pour tout ω .

Bien entendu, le problème mathématique de minimisation (ou maximisation) de $F(x, \omega)$ sous des contraintes $x \in X$ dépend des réalisations de ω et n'a donc pas réellement de sens. Résoudre ce problème fournit autant de solutions possibles que nous disposons de réalisations de ω sans vraiment pouvoir affirmer quelle est la meilleure de toutes. Une façon de régler le problème est d'optimiser la fonction objectif en *espérance* (ou en moyenne si nous considérons que toutes les réalisations de ω sont équiprobables). Le problème d'optimisation stochastique général s'écrit donc :

$$\min_{x \in X} f(x) := \mathbb{E}[F(x, \omega)]. \quad (2.1)$$

Cette formulation d'un problème stochastique suppose implicitement que l'espérance est calculée en connaissance de la distribution de probabilité P et en supposant que l'opérateur

$$\mathbb{E}[F(x, \omega)] = \int_{\Omega} F(x, \omega) dP(\omega)$$

est bien défini.

Un problème de ce type est appelé problème *statique* (ou 1-étape ou “Here and Now”). Cela signifie que, bien que le problème soit défini sur plusieurs périodes de temps, toutes les décisions ne sont prises qu’en une seule fois, et ce dès l’instant initial t_0 de la période d’étude. Il n’y a donc aucun moyen de corriger les décisions prises au cours du temps en fonction des réalisations successives de l’aléa ω . Nous disons également que nous sommes en *boucle ouverte* pour signifier que, en tant que fonction de l’aléa ω , la commande optimale x est une fonction constante.

Nous pouvons souligner deux difficultés à cette approche statique :

- Il peut être assez difficile en pratique de posséder un modèle probabiliste de l’aléa. En effet, certains phénomènes aléatoires peuvent être relativement complexes, voire impossibles à modéliser. Nous pouvons parfois réussir à caler des modèles statistiques sur des historiques sans réelle garantie de qualité. L’hypothèse de connaissance parfaite de P est donc en fait une hypothèse assez forte ;
- D’un point de vue théorique, nous pouvons calculer, en tout point x par lequel cheminera l’algorithme d’optimisation choisi pour résoudre le problème, la valeur $f(x)$ du critère et son sous-gradient $\partial f(x)$, si bien que le problème d’optimisation stochastique (2.1) se ramène au problème déterministe $\min_{x \in X} f(x)$: les méthodes d’optimisation convexe générale (Faisceaux, ACCPM,...) s’y appliquent alors sans difficulté particulière. Cependant, d’un point de vue pratique, chaque évaluation de $f(x)$ ou de $\partial f(x)$ nécessite un calcul d’espérance, ce qui peut s’avérer très consommateur en temps de calcul sur un ordinateur si la variable aléatoire ω prend ses valeurs dans un espace multidimensionnel. Typiquement, pour une distribution continue, le calcul d’intégrales multiples nécessaire pour le calcul de l’espérance est impossible pour une bonne précision quand le nombre de dimensions excède 4 [70].

2.1.1 Le gradient stochastique

Une approche mieux adaptée que les approches classiques de décomposition à ce type de problème est celle du gradient stochastique ([22, 35, 106]), dont le principe est de faire évoluer simultanément le calcul de l’espérance et la méthode de gradient. L’idée consiste à tirer une suite $(\omega^1, \dots, \omega^k, \dots)$ de réalisations indépendantes de la variable ω suivant sa loi de probabilité, et à faire évoluer la variable x à chaque fois qu’une nouvelle valeur de l’aléa est disponible en effectuant un pas de gradient sur la fonction F en fixant ω à cette valeur ; l’évolution de x doit être suffisamment lente pour que le phénomène de moyenne lié à l’espérance se fasse au cours des itérations. Le grand avantage de cette méthode est donc le calcul totalement implicite de l’espérance. Plus précisément, la forme générale de l’algorithme de gradient stochastique est la suivante :

- nous nous donnons $x^0 \in X$ pour initialiser l’algorithme ;
- à l’itération k , nous effectuons un tirage ω^{k+1} de l’aléa ω suivant sa loi, et nous effectuons un pas de gradient en x : $x^{k+1} = \text{proj}_X(x^k - \epsilon^k F'_x(x^k, \omega^{k+1}))$, où ϵ^k est le terme d’une suite de réels positifs qui converge “lentement” vers zéro.

Cet algorithme, ainsi que ses variantes, est ancien. Nous pouvons citer les initiateurs Robbins et Monro [88], Kiefer et Wolfowitz [56]. Le premier résultat de complexité est dû à Nemirovski

et Yudin [68]. Polyak ([78, 79]), Polyak et Tsypkin [80] donnent des résultats de convergence globale (sous des hypothèses de convexité) ainsi que des études de vitesse de convergence et diverses variantes de cet algorithme. Dans Kushner et Clark [58], nous trouverons la méthode de l'équation différentielle moyenne qui permet l'étude de la convergence locale de cet algorithme et de ses variantes. Cette approche est très utile dans le cas non convexe. Comme résultat majeur, nous pouvons citer les travaux de Nesterov et Vial [69] qui, en introduisant la notion de ϵ -optimalité avec un niveau de confiance, montrent l'existence d'une solution au problème général et donnent des bornes de complexité (sur le nombre d'itérations de calculs). L'article [54] compare notamment les approches par scénario et de gradient stochastique. Les auteurs montrent qu'avec des pas bien choisis, les résultats obtenus par gradient stochastique peuvent être aussi satisfaisants que pour des approches par scénarios, bien que nous utilisons des calculs de gradient. Enfin, nous pouvons citer quelques applications à la gestion de production électrique ([31, 43, 2]).

2.1.2 L'approche par scénarios

La principale difficulté de ce problème est le calcul de l'espérance. Une fois ce problème résolu, le problème devient un problème classique d'optimisation. Une technique classique pour approximer l'espérance est le calcul par scénarios. Nous supposons posséder une distribution finie des aléas et nous écrivons l'espérance sous la forme d'une moyenne pondérée. Nous reviendrons longuement sur ce type d'approximation en section 2.2.

2.1.3 L'aléa dans les contraintes

Dans le modèle (2.1), nous supposons posséder une description explicite de l'ensemble de faisabilité X . Par exemple, l'ensemble X peut être écrit sous une forme standard en programmation mathématique :

$$X := \{x \in X_0 : g_i(x) \leq 0, i = 1, \dots, m\},$$

où X_0 est un sous-ensemble convexe de χ et les $g_i(x)$ sont des fonctions à valeurs réelles. Mais si les contraintes sont, à l'image de la fonction objectif, soumises à l'aléa, il nous faut spécifier ce que signifie d'être réalisable. Une valeur particulière de x peut vérifier les contraintes pour certaines valeurs de ω et les violer complètement pour d'autres ω .

Une première approche peut consister à vouloir satisfaire les contraintes pour toute réalisation de l'aléa dans Ω . Une telle approche est clairement irréaliste et conservatrice. Plusieurs approches peuvent être utilisées pour introduire un sens à la réalisabilité de ces contraintes.

Une d'entre elles est de considérer l'espérance mathématique :

$$g_i(x) := \mathbb{E}[G_i(x, \omega)], i = 1, \dots, m.$$

Un telle approche est souvent utilisée en pratique quand le problème possède en réalité plusieurs objectifs et que nous voulons en substituer quelques uns à des contraintes, à l'image d'une contrainte de risque. Pour résoudre ces contraintes, nous utilisons en général les approches par scénarios.

Une autre façon de définir l'ensemble de faisabilité est d'utiliser des contraintes de satisfaction en probabilité :

$$P[G_i(x, \omega) \leq 0] \geq 1 - \alpha, \quad i = 1, \dots, m. \quad (2.2)$$

Les contraintes en probabilité ont été introduites par Charnes et al. en 1958 [24] sous la forme individuelles, c'est à dire en considérant la probabilité de chaque contrainte. Elles ont été introduites sous la forme jointe (la probabilité couple plusieurs contraintes) par Miller et Wagner [64] mais en considérant des aléas indépendants situés uniquement dans le second membre. C'est Prékopa [81, 83] qui initia le modèle plus général où les contraintes en probabilité sont jointes et les variables aléatoires dépendantes les unes des autres.

Même si l'écriture paraît attrayante, ces contraintes sont très difficiles à gérer généralement. Les principaux résultats sont, en supposant que $G_i(x, \omega) = Ax - \omega$:

- Dans le cas d'une contrainte individuelle, avec un aléa ω mono-dimensionnel et séparé des variables, nous pouvons transformer la contrainte en probabilité en une inégalité linéaire par un simple calcul de quantile [50, 20] ;
- Pour le cas où la matrice A est déterministe et ω a une densité de probabilité multidimensionnelle log-concave, Prékopa [82] a montré que l'ensemble de faisabilité est convexe.

D'autres cas standard sont répertoriés dans [34]. Par ailleurs, ces contraintes étant de nouveau en vogue, de nouveaux résultats de convexité apparaissent dans la littérature (par exemple [51]). Cependant, en dehors de ces cas particuliers, l'ensemble de faisabilité est généralement non convexe, voire non connexe [50, 49]. Certains exemples élémentaires illustrent cela au chapitre 3.

Ces contraintes étant difficiles à prendre en compte dans un modèle en raison de leurs propriétés mathématiques, il est plus facile d'essayer de les transformer en contraintes convexes. Nous pouvons citer trois approches principales :

- Approche par scénarios : en supposant que la distribution de probabilité soit bien connue et que nous puissions générer des scénarios, nous remplaçons la contrainte en probabilité par un ensemble de contraintes à satisfaire sur chaque scénario. Le nombre de scénarios à considérer dépend des niveaux de probabilité et de précision souhaités. Nous pouvons citer sur ces approches les travaux de Shapiro et Nemirovski [66] et Bonami et Lejeune [20] ;
- Approximation de Bernstein : en supposant que les contraintes soient affines dans les perturbations et que les éléments du vecteur aléatoire sont indépendants entre eux, la contrainte en probabilité peut être approximée par une contrainte convexe à l'aide de la formule de Bernstein [67] ;

- L’optimisation robuste : si nous sommes capables de construire un ensemble d’incertitude qui contienne les $1 - \alpha$ éléments les plus probables, les commandes obtenues par optimisation robuste alors vérifieront la contrainte en probabilité 2.2. Nous reviendrons sur ce lien entre optimisation robuste et contrainte en probabilité dans le chapitre 3.

2.2 Les problèmes à deux étapes

Les problèmes de décision peuvent être plus complexes que dans le cas statique. Imaginons le cas d’un magasin. Le responsable doit prendre une décision initiale à t_0 pour remplir son stock en fonction de prévisions de demande. Au bout d’une certaine période, le responsable a, à la vue de la demande réalisée, la possibilité de commander ou non des articles supplémentaires. Ce problème est dit *à deux étapes* car deux types de décisions doivent être prises : une décision initiale et une décision dite de recours après observation de la réalisation de l’aléa. L’expression “à deux étapes” ne signifie donc pas que le problème ne comporte que deux pas de temps mais plutôt deux étapes de décision distinctes.

Le problème stochastique linéaire à deux étapes peut s’écrire de façon générale par

$$\min_{x \in X} \{g(x) := c^T x + \mathbb{E}[Q(x, \xi(\omega))]\}, \quad (2.3)$$

où $Q(x, \xi) = \min_{y \in Y} \{q^T y \mid Tx + Wy \leq h\}$ est la valeur optimale du problème de seconde étape.

Ici, $x \in \mathbb{R}^n$ est la décision de première étape, X un polyèdre défini comme un nombre fini de contraintes linéaires, $y \in \mathbb{R}^m$ le vecteur de décisions de seconde étape défini sur Y et $\xi = (q, T, W, h)$ le vecteur contenant toutes les données du problème de seconde étape. En réalité, ξ devrait s’écrire $\xi(\omega)$ et donc peut être vu comme un vecteur aléatoire. L’espérance de Q est calculée sous l’hypothèse de connaissance de la distribution de probabilité de $\xi(\omega)$. Les matrices T et W sont appelées respectivement matrices *technologiques* et de *recours*.

Avec ces formulations, en première étape, nous prenons une décision ferme x de coût $c^T x$ avant que la réalisation de l’aléa ξ soit révélée. Dès que cette information est connue, en seconde étape, nous optimisons notre recours y de coût $q^T y$ pour que Wy compense une possible inconsistance du système $Tx \leq h$. Ce concept de problème stochastique à deux étapes a été introduit par Beale [4] et Dantzig [26] et discuté dans [17, 55, 34, 92, 101].

Si X et Y sont des polyèdres, la fonction $Q(\cdot, \xi)$ est convexe, impliquant la convexité du problème à deux étapes. L’hypothèse de convexité est donc fondamentale.

Nous pouvons introduire quelques définitions supplémentaires :

- Nous parlons de matrice de recours fixe si la matrice W n’est pas aléatoire ;
- En particulier, nous parlons de recours fixes complets si la matrice fixée, de taille $m \times \bar{n}$, W est surjective c’est-à-dire satisfait

$$\{z \mid z = Wy, \forall y\} = \mathbb{R}^m.$$

Ceci implique que, quelque soit la décision de première étape x et la réalisation ξ , le problème de seconde étape $Q(x, \xi) = \min\{q^T y \mid Wy = h(\xi) - T(\xi)x\}$ est toujours réalisable. Autrement dit, nous pouvons toujours trouver un recours réalisable. Cette définition demeure vraie même pour une matrice W non fixée ;

- Enfin, un cas particulier de matrice de recours fixe est le recours simple où $(T, W) = (Id, -Id)$.

Il peut arriver que l'information observée soit incomplète. Dans ce cas, il est nécessaire pour la faisabilité des recours qu'ils soient mesurables par rapport aux informations observables [104].

Une approche standard pour résoudre les problèmes à deux étapes est par construction de scénarios : nous supposons que le vecteur aléatoire ξ a un nombre fini de réalisations possibles ξ_1, \dots, ξ_K de probabilités respectives p_1, \dots, p_K . L'espérance peut ainsi être écrite comme la somme finie

$$\mathbb{E}[Q(x, \xi)] = \sum_{k=1}^K p_k Q(x, \xi_k).$$

En attribuant une décision de recours $y_k = y(\xi_k)$ à chaque scénario et en se ramenant au cas linéaire, nous pouvons écrire le problème de recours comme un gros problème d'optimisation linéaire :

$$\min_{x, y_1, \dots, y_K} \quad c^T x + \sum_{k=1}^K p_k \langle q_k, y_k \rangle \tag{2.4}$$

$$\text{s.t.} \quad x \in X, \ y_k \in Y, \ T_k x + W_k y_k \leq h_k, \ k = 1, \dots, K. \tag{2.5}$$

En résolvant ce problème, nous obtenons une solution optimale \bar{x} de première étape et une série de recours \bar{y}_k , pour chaque scénario correspondant à la réalisation $\xi = \xi_k$. Au cours du temps, de nombreux efforts ont été effectués pour développer des algorithmes exploitant les structures du problème (2.4) (Voir [91] pour une vue d'ensemble sur les décompositions de problèmes de ce type).

Discutons de deux conditions nécessaires pour la résolution du problème à deux étapes :

- $Q(x, \xi)$ est définie comme la valeur optimale du problème de seconde étape. Il peut arriver que, pour un contrôle réalisable $x \in X$ et un scénario ξ_k , le problème n'est pas borné inférieurement, i.e. $Q(x, \xi_k) < -\infty$. Cela veut dire que pour un contrôle réalisable x , nous pouvons améliorer le coût indéfiniment avec une probabilité positive, situation à éviter ;
- un problème plus sérieux est que, pour un $x \in X$ et un scénario ξ_k , le système $T_k x + W_k y \leq h_k$ est incompatible, i.e. que le problème de seconde étape est infaisable. Nous imposons donc $Q(x, \xi_k) = +\infty$ pour un tel contrôle x , ce qui l'élimine. En clair, nous imposons une pénalité

2.2. LES PROBLÈMES À DEUX ÉTAPES

de coût infini si le problème de seconde étape n'est pas réalisable, ce qui oblige un ajustement de la décision de première étape. Si une telle infaisabilité n'intervient pas, nous disons que le problème possède des recours complets. Il est toujours possible de rendre le problème de deuxième étape réalisable, c'est-à-dire relaxer le problème avec recours en un problème avec recours complet, en étendant le recours y à (y, t) , en pénalisant t et en relaxant la contrainte de seconde période. Pour cela, nous écrivons le problème :

$$\min_{y,t} q^T y + \gamma t \text{ s.t. } Tx + Wy - te \leq h, \quad t \geq 0,$$

où $\gamma > 0$ est une constante choisie et e un vecteur dont toutes les composantes sont 1. Dans cette formulation, la pénalité d'une possible infaisabilité est contrôlée par le paramètre γ . L'approche standard consiste à introduire une coupe de réalisabilité [91].

La génération de scénarios est une étape nécessaire mais difficile. Le générateur doit inclure une modélisation (loi,...) la plus riche possible pour que les scénarios représentent au mieux les phénomènes aléatoires physiques à prendre en compte. La qualité des scénarios est, en effet, primordiale puisque même un très grand nombre de scénarios ne permettrait pas de résoudre efficacement ce problème si les scénarios ne représentent pas les phénomènes aléatoires de façon correcte.

Le nombre de scénarios est aussi une question importante. Il est primordial de trouver le compromis adéquat entre générer un très grand nombre de scénarios et donc résoudre un problème de taille très grande et générer un faible nombre de scénario, nombre qui pourrait s'avérer insuffisant pour une résolution correcte du problème. Nous pouvons prendre un petit exemple pour illustrer cela. Supposons que les composantes de $\xi \in \mathbb{R}^d$ soient indépendantes entre elles et peuvent ne prendre que r différentes valeurs. Le nombre total de scénarios est $K = r^d$. K grandit donc exponentiellement quand r ou d croît. Si $r = 4$ et $d = 20$, nous nous retrouvons avec le nombre astronomique de scénarios $4^{20} \approx 10^{12}$. Il semble donc difficile de calculer exactement l'espérance avec une précision élevée. Une approche possible pour estimer cette espérance, tout en conciliant taille et précision, est l'application des techniques d'échantillonnage Monte-Carlo. Approximer l'espérance correspond à trouver une borne N sur le nombre de scénarios nécessaires pour garantir avec une bonne probabilité que l'espérance est estimée à une erreur ϵ près.

L'objectif d'une approche Monte-Carlo est donc de réduire l'ensemble de scénarios disponible en un ensemble de scénarios de taille gérable numériquement mais suffisant pour obtenir une solution satisfaisante. Nous supposons que l'ensemble de scénarios initial est très grand voire infini. En se basant sur l'échantillon des ξ , nous pouvons ainsi estimer l'espérance $q(x) = \mathbb{E}[Q(x, \xi)]$ par la moyenne

$$\hat{q}_K(x) = \frac{1}{K} \sum_{k=1}^K Q(x, \xi^k),$$

et par conséquent l'espérance du problème (2.3) :

$$\min_{x \in X} \left\{ \hat{g}_K(x) := c^T x + \frac{1}{K} \sum_{k=1}^K Q(x, \xi^k) \right\}.$$

Cette idée simple est connue dans la littérature récente sous le nom de *sample average approximation method* (SAA) [97, 60, 19, 98, 100, 92, 101]. SAA n'est pas un algorithme mais juste une reformulation du problème. Nous devons toujours résoudre le problème par la procédure numérique adéquate.

Nous pouvons donc retourner à la question : “quel nombre K de scénarios devons nous prendre pour le problème SAA soit une bonne approximation de 2.3?”. Les approches Monte-Carlo sont connues pour leur faible vitesse de convergence. Pour un $x \in X$ fixé et un échantillon iid (chaque scénario est indépendant), la variance de l'échantillon moyen $\hat{q}_K(x)$ vaut :

$$\text{Var}(\hat{q}_K(x)) = \frac{\text{Var}(Q(x, \xi))}{K}.$$

Ceci implique que l'échantillon moyen converge vers l'espérance correspondante à un taux de $O_p(K^{-1/2})$. Cela signifie que pour l'amélioration de la précision d'une décimale, nous devons augmenter la taille de l'échantillon d'un facteur 100. Il n'est donc pas avantageux d'utiliser des approches Monte-Carlo quand le nombre d de variables aléatoires est petit. Nous préférons ainsi discrétiser directement les aléas ou utiliser des approches Quasi Monte-Carlo [70]. Nous pouvons trouver dans [97] des bornes sur le nombre de scénarios à prendre en compte pour un certain niveau de précision. Il faut ajouter que les temps de calculs sont toutefois dépendants de la dimension de l'aléa. Enfin, précisons qu'il est toujours possible d'estimer l'espérance par Monte-Carlo en dimension importante si nous n'optimisons pas sur tous les scénarios nécessaires.

L'avantage avec les techniques Monte-Carlo est que la précision de l'approximation obtenue ne dépend pas de la taille de l'échantillon mais seulement de la variance de $Q(x, \xi)$. Il a été montré théoriquement et vérifié expérimentalement [97] que la méthode SAA aide à la résolution du problème à deux étapes avec une précision raisonnable (1% ou 2%) sous les conditions suivantes :

1. il est possible de générer un échantillon de réalisations du vecteur aléatoire ξ ,
2. pour des tailles raisonnables d'échantillons, il est possible de calculer efficacement le problème SAA obtenu,
3. le problème à deux étapes a des recours complets,
4. la variabilité de la fonction de seconde étape n'est pas trop grande.

Ce type de résultats peut être retrouvé dans [60, 61, 77, 93, 105]. Il a été aussi démontré que des techniques Quasi-Monte Carlo peuvent améliorer la précision de SAA significativement (voir [70] pour une discussion général sur Quasi-Monte Carlo et [27, 57, 73] pour des applications à la programmation stochastique). Enfin, nous pourrions croire que cette difficulté à résoudre les problèmes à deux étapes provient de l'approche par scénarios. Mais Dyer et Stougie [33] ont montré que les problèmes stochastiques linéaires à deux étapes sont NP-difficiles, même avec des recours complets et des distributions discrètes simples à décrire.

Le problème à deux étapes défini s'écrit également :

$$\begin{aligned} & \min_{x \in X} \{ \mathbb{E}[c^T x + \min_{y \in Y} \{q^T y : Wy \leq h - Tx\}] \} \\ \Leftrightarrow & \min_{x \in X} \{ \mathbb{E}[\phi(x, \xi)] \} \end{aligned}$$

Par cette écriture, nous nous rendons compte que le problème à deux étapes peut se ramener à un problème statique. Si le problème est convexe (donc si ϕ est convexe en x), les méthodes générales d'optimisation convexe peuvent s'appliquer. De même, les approches par scénarios présentées dans cette section s'étendent au cas statique.

2.3 Les problèmes à plusieurs étapes

Dans cette section, nous généralisons à T étapes le problème à deux étapes. Les décisions sont donc prises en T fois, à partir de l'information disponible en chaque $t = 1, \dots, T$. Ce problème s'écrit sous sa forme générale :

$$\min_{x_1 \in \mathcal{G}_1} F_1(x_1) + \mathbb{E} \left[\inf_{x_2 \in \mathcal{G}_2(x_1, \xi_2)} F_2(x_2, \xi_2) + \mathbb{E} \left[\dots + \mathbb{E} \left[\inf_{x_T \in \mathcal{G}_T(x_{T-1}, \xi_T)} F_T(x_T, \xi_T) \right] \right] \right], \quad (2.6)$$

où les $x_t \in \mathbb{R}^{n_t}$ sont les variables de décision, $F_t : \mathbb{R}^{n_t} \times \mathbb{R}^{d_t} \rightarrow \mathbb{R}$ sont des fonctions continues et \mathcal{G}_t , $t = 2, \dots, T$ sont les ensembles de faisabilité fermés de chaque étape t . F_1 et \mathcal{G}_1 sont déterministes.

Dans le cas multidimensionnel, chaque décision est prise après la réalisation de l'aléa. Chaque décision est donc prise en connaissance de l'histoire du processus mais également en toute ignorance de son futur. Ce principe est le principe de *non-anticipativité*. Nous introduirons pour la suite la filtration ξ , représentant l'histoire du processus, à l'aide de la notation $\xi_{[t]} := (\xi_1, \dots, \xi_t)$.

Contrairement au problème à deux étapes, la difficulté principale provient du fait que nous ne pouvons plus nous ramener à un problème statique. Les problèmes de minimisation imbriqués empêchent en effet d'avoir une fonction ne dépendant uniquement que des variables de première étape.

Nous pouvons retrouver les résultats sur ces problèmes à plusieurs étapes notamment dans [101, 92].

2.3.1 La programmation stochastique sur arbre

Une approche standard pour résoudre les problèmes stochastiques à plusieurs étapes est la programmation stochastique basée sur une discrétisation de l'aléa sous une forme d'arbre (Fig. 6.2) : le problème original est résolu en espérance sur une distribution discrète et connue de l'aléa. Les difficultés de cette approche peuvent être résumées en deux questions : comment construire

l'arbre et comment optimiser dessus ?

Un premier point sur cette approche est que la distribution de probabilité est fixée. Cela signifie que les décisions ne peuvent pas influencer sur la distribution de l'aléa ce qui est une certaine limitation (par exemple, décider de lancer une campagne publicitaire peut influencer fortement la demande). C'est pourquoi l'arbre sur lequel nous optimisons est construit à l'instant initial et n'est plus modifié.

La construction s'effectue comme suit : à $t = 1$, l'aléa ne prend qu'une unique valeur ξ_1 . A $t = 2$, chaque réalisation ξ_2^i est représentée par un nœud i . A $t = 3$, nous relierons ξ_2^i à toutes les réalisations de ξ_3 qui peuvent se réaliser à la suite de ξ_2^i . Formellement, nous relierons le nœud ξ_t^i à ses nœuds fils ξ_{t+1}^{ij} par un arc, correspondant à une probabilité de réalisation $p_{ij} > 0$ avec $\sum_j p_{ij} = 1$. Généralement, de tels arbres sont construits à partir de scénarios et nous utilisons des techniques d'agrégation [76, 48] : à chaque instant, nous agrégeons en un unique nœud tous les scénarios proches au sens d'une certaine distance à définir. Si l'arbre d'aléas représente la réalité et est construit sur la base d'un nombre raisonnable de scénarios, la complexité numérique de résolution de ce problème est certes élevée mais pas plus difficile qu'un problème à deux étapes avec le même nombre de scénarios. Par contre, si le nombre de scénarios est très élevé, les problèmes à deux ou plusieurs étapes ne peuvent être résolus. Cependant, nous avons pu voir que certaines classes de problèmes à deux étapes pouvaient être résolus à l'aide de techniques d'échantillonnage Monte Carlo. Ces approches SAA peuvent également être appliquées pour les problèmes à plusieurs étapes : nous générons N_1 réalisations de ξ_2 qui représentent les nœuds de probabilités respectives $1/N_1$. Puis nous générons N_2 scénarios pour $t = 3$ en prenant en compte les distributions conditionnelles de ξ_3 sachant la réalisation ξ_2^i . De la sorte, nous construisons un arbre en générant $N = \prod_{t=1}^{T-1} N_t$ scénarios. Cette approche de construction des arbres est un échantillonnage conditionnel. Des études sur la stabilité de ces approches ont été effectuées [85, 36, 86, 87]. Le nombre de scénarios grandit donc exponentiellement avec le nombre de périodes de temps.

Une fois l'arbre construit, nous pouvons formuler l'équivalent déterministe en attribuant une commande à chaque nœud de l'arbre, c'est à dire une commande à chaque réalisation de l'aléa. Si le problème stochastique original est linéaire, l'équivalent déterministe obtenu sera toujours un problème d'optimisation linéaire mais d'une taille bien plus importante. Dans [98, 96, 99, 97], les auteurs avancent que "les problèmes à plusieurs étapes, même linéaires, avec des recours complets sont généralement numériquement intraitables même pour une précision faible". Ce constat n'empêche pas que "certains problèmes puissent être résolus efficacement". Dyer et Stougie avancent pour leur part que "la programmation stochastique multi-étapes est au moins aussi complexe que les problèmes combinatoires difficiles". Pour remédier à cela, il existe des techniques de réduction d'arbre [47, 32] qui permettent de réduire la taille de l'arbre tout en gardant l'information principale. Utiliser la programmation stochastique consiste donc à faire un arbitrage entre les temps de calculs et la précision de la solution.

Cette approche est très répandue et est particulièrement développée dans l'industrie électrique [46, 44, 71, 45, 41, 29, 38, 23, 42, 102, 72, 21].

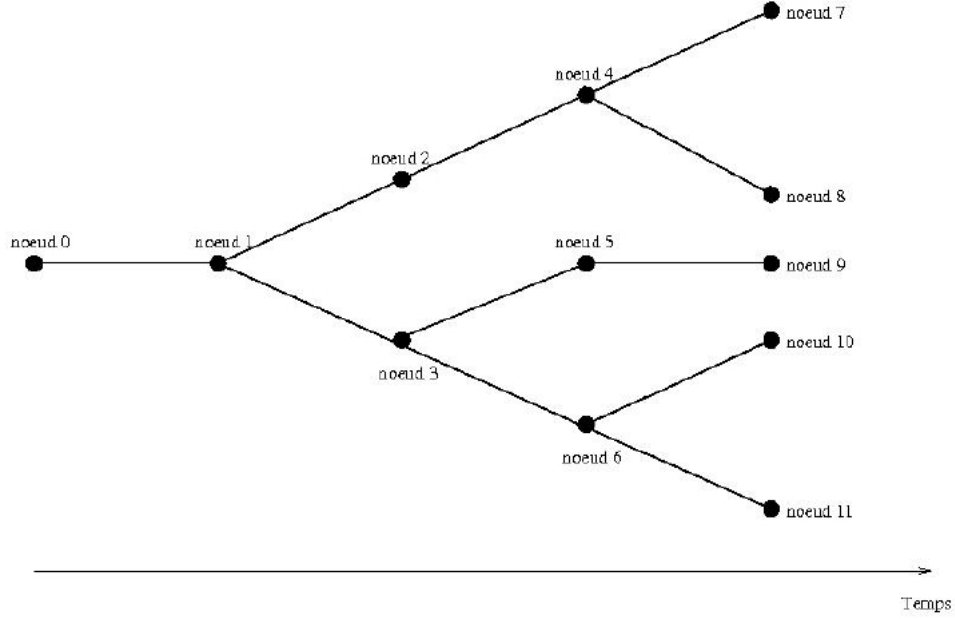


FIG. 2.1 – Représentation de l'aléa par un arbre

2.3.2 La programmation dynamique et ses extensions

Une autre façon de résoudre ce problème est la programmation dynamique introduite par Bellman [6, 7, 8]. Pour cela, considérons le problème de dernière étape

$$\min_{x_T \in \mathcal{G}_T(x_{T-1}, \xi_T)} F_T(x_T, \xi_T).$$

La valeur optimale de ce problème, notée $V_T(x_{T-1}, \xi_T)$, dépend du vecteur de décision x_{T-1} et des données ξ_T . En introduisant l'histoire $\xi_{[t]}$ du processus ξ , nous pouvons écrire le problème général pour les pas de temps $t = 2$ à $T - 1$:

$$\begin{aligned} \min_{x_t} \quad & F_t(x_t, \xi_t) + \mathbb{E}[V_{t+1}(x_{t+1}, \xi_{[t+1]}) \mid \xi_{[t]}] \\ \text{s.t.} \quad & x_t \in \mathcal{G}_t(x_{t-1}, \xi_t), \end{aligned}$$

où $\mathbb{E}[\cdot \mid \xi_{[t]}]$ est l'espérance conditionnelle à l'histoire de ξ . La valeur optimale de chacun de ces problèmes, notée $V_t(x_{t-1}, \xi_{[t]})$ dépend de la décision prise en $t - 1$ et de toutes les réalisations passées $\xi_{[t]}$. L'idée de la programmation dynamique est de calculer récursivement ces fonctions valeurs $V_t(x_{t-1}, \xi_{[t]})$ en partant de T jusqu'à $t = 1$. C'est la phase rétrograde de la programmation dynamique (ou "backward").

Cependant, nous pouvons distinguer deux difficultés majeures dans cette formulation de la programmation dynamique :

- Calculer les fonctions V_t peut être vraiment difficile voire impossible à cause de l'espérance conditionnelle. Une façon de simplifier ce problème est de considérer que le processus aléatoire est Markovien. La fonction valeur n'est donc plus fonction de $\xi_{[t]}$ mais uniquement de ξ_t . La fonction valeur en t s'écrit ainsi $V_t(x_{t-1}, \xi_t)$ et est désormais plus aisée à calculer ;
- Pour calculer la fonction valeur en chaque date t , la fonction valeur doit être parfaitement connue pour toutes les valeurs possibles de l'espace d'état sous-jacent. Sauf à avoir une formule analytique explicite pour les fonctions valeur, cela n'est pas réalisable. En programmation dynamique classique, la méthodologie consiste à discrétiser l'espace des états. La fonction valeur est ainsi calculée en chaque point de ce maillage grâce à la dynamique. L'approximation de la fonction de recours est ensuite obtenue en interpolant l'ensemble de ces valeurs sur un pas de temps donné. Si nous considérons n états discrétisés chacun en d points, nous devons calculer d^n transitions entre chaque point du maillage pour chaque pas de temps. Avec une description un minimum fine de l'état et sur une période de temps importante, nous atteignons assez rapidement des temps de calculs inconsidérables. La conséquence est que ce calcul peut devenir irréaliste si le nombre d'états est trop important. C'est la "malédiction de la dimension".

Il existe une grande littérature sur le programmation dynamique. Nous citerons notamment les travaux de Bertsekas et al. [13, 15], Denardo [28], Hinderer [52], Nemhauser [65], Puterman [84], Ross [89] ou Sennott [95].

Pour palier aux problèmes de dimension et souvent au détriment de l'optimalité, un grand nombre d'approximations ont été mises en place. Nous pouvons citer de façon totalement exhaustive : des approches basées sur la discrétisation des états [108, 12, 59], des approches utilisant des problèmes de petite taille et des espaces d'état finis pour résoudre des problèmes où l'espace d'états est infini [37, 94, 30, 107], des approches d'aggrégation d'états [5, 14, 25, 63], ...

Un autre type d'approximation est l'approximation des fonctions valeur. C'est le principe de base de la *programmation dynamique duale*. Cet algorithme est très utilisé en gestion de la production et en particulier pour la gestion hydraulique [18, 40, 62, 74, 75, 90].

Elle contourne le problème de la dimension lié à la discrétisation de l'espace des états en exploitant la structure de la fonction valeur. En effet, nous pouvons montrer que sous certaines hypothèses (par exemple, la linéarité), la fonction valeur est convexe. Elle peut donc être représentée comme supportée par un nombre fini d'hyperplans. La programmation dynamique duale exploite ce résultat afin de construire itérativement des hyperplans de support pour représenter de mieux en mieux la fonction valeur. Par conséquent, il n'y a plus de discrétisation préalable de l'espace des états et la malédiction de la dimension peut être contournée.

Chaque itération de l'algorithme comporte, à l'image de la programmation dynamique, deux phases : une phase rétrograde d'optimisation qui permet d'améliorer l'approximation des fonctions valeur en rajoutant des hyperplans et une phase de simulation qui permet d'obtenir une borne supérieure sur le coût. Les problèmes résolus à chaque itération sont des problèmes de petite taille, simples à résoudre.

Toutefois, cette approche comporte quelques limites : pour une itération fixée, nous résolvons à chaque instant le problème d'optimisation de la fonction valeur, problème qui dépend du pas de temps suivant. L'erreur d'approximation de la fonction valeur introduite en T se répercute ainsi au fil du temps et même grandit au fil du temps. Donc plus le nombre de pas de temps est important, plus l'approximation des fonctions valeur sur les premiers pas de temps sera grossière. Un nombre de pas de temps trop important est donc problématique, d'autant plus que les temps de calculs en sont fortement influencés. D'autre part, cette approche est heuristique et beaucoup de paramètres doivent être réglés manuellement (comme le choix des points candidats servant de support aux hyperplans). Enfin, les critères de convergence ne renseignent pas sur la distribution des coûts obtenus. En effet, le test d'arrêt ne se concentre que sur l'espérance du coût. Dans une optique de gestion des risques sur des scénarios extrêmes, les stratégies de la SDDP pourraient être très éloignées des stratégies optimales.

2.3.3 Les autres approches

Nous avons présenté des approches dans la cas statique ou à deux étapes. Pouvons-nous les généraliser ?

Pour appliquer l'algorithme du gradient stochastique, l'hypothèse de séparabilité dans le temps est fondamentale. Dans ce cas de boucle fermée, il est donc pas possible d'appliquer ce type d'algorithme. Cependant des premières pistes d'amélioration apparaissent dans la littérature [3].

En fait, à la vue de la difficulté des problèmes à plusieurs étapes, des techniques sous-optimales ont été introduites. C'est le cas des règles de décision. Ce concept de règle de décision a été introduit dans les années 60 dans le contexte du contrôle optimal. Il a été développé pour résoudre des problèmes de planning de production [53] avec un objectif quadratique. Depuis, ces règles de décision linéaires ont été utilisées avec de la programmation stochastique et des contraintes en probabilité [39] mais oubliées depuis. Elles sont réapparues récemment avec les travaux en optimisation robuste ([10, 9, 11, 1, 16]). Il s'agit de remplacer les commandes par des règles de décision linéaires en l'observation de l'aléa puis de construire l'équivalent robuste par rapport à ces aléas (voir chapitre 3). Au sujet des contraintes en probabilité, il semble que ces règles de décision soient à l'heure actuelle la seule façon d'introduire de la dynamique. Mais ceci n'est possible que sous de sévères hypothèses comme la séparabilité de la variable aléatoire. En programmation stochastique, nous pouvons souligner l'introduction de règles de décision constantes par morceaux ([103] et chapitre 6), heuristique qui sera présentée plus longuement dans cette thèse.

2.4 Conclusions

Nous venons de faire un rapide survol de l'optimisation dans l'incertain. Les problèmes d'optimisation qui concernent cette thèse sont clairement multidimensionnels. La programmation dynamique, à cause de la malédiction de la dimension, et la programmation stochastique, à

cause de la difficulté numérique inhérente, ne semblent pas satisfaisantes pour résoudre des problèmes à grande dimension. Or, le cadre industriel de cette thèse réclame la nécessité de proposer des approches qui soient cohérentes mais surtout qui répondent à des besoins particuliers comme des temps de calculs raisonnables. D'autre part, les commandes que nous pouvons chercher à obtenir comme, par exemple, des turbinages de centrales hydrauliques, se doivent de respecter au maximum les contraintes pour que les exploitants n'aient pas à faire des corrections incessantes et importantes qui, finalement, désoptimisent le système. A la vue de nos objectifs et notamment la volonté d'avoir une approche opérationnelle, deux approches semblent particulièrement intéressantes. La première est l'optimisation robuste. L'optimisation robuste permet de s'assurer des contrôles réalisables pour toute réalisation d'un aléa dans un ensemble contenant ses réalisations les plus probables. La particularité principale de cette approche est la volonté affichée de conserver une complexité de l'équivalent déterministe équivalente à celle du problème déterministe original. Nous détaillerons et appliquerons donc cette approche notamment au sein des chapitres 3 et 4. La seconde est la Programmation Stochastique avec Règles de Décision Constantes par Morceaux. Il s'agit d'une heuristique basée sur la programmation stochastique classique sur arbre. Le grand avantage est qu'elle permet de contrôler la taille de l'arbre tout en obtenant des résultats intéressants. Nous présenterons cette approche et l'appliquerons au sein du chapitre 6.

Bibliographie

- [1] E. Adida and G. Perakis. A Robust Optimization Approach to Dynamic Pricing and Inventory Control with no Backorders. *Mathematical Programming*, 107(1-2) :97 – 129, 2006.
- [2] M. Balducci, G. Cohen, J.C. Dodu, M. Goursat, A. Hertz, J.P. Quadrat, and M. Viot. Comparaison de Trois Méthodes d'Optimisation des Investissements dans un Réseau de Transport d'Electricité. In *Conference IFAC*, 1981.
- [3] K. Barty, J.S. Roy, and C. Strugarek. A Stochastic Gradient Type Algorithm for Closed Loop Problems. *Stochastic Programming E-print Series*, 2005-14, 2005.
- [4] E.M.L. Beale. On Minimizing a Convex Function Subject to Linear Inequalities. *J.R. Stat. Soc.*, 17 :173–184, 1955.
- [5] E.M. Bean, J.R. Birge, and R.L. Smith. Aggregation in Dynamic Programming. *Operations Research*, 35 :215–220, 1987.
- [6] R.E. Bellman. Dynamic Programming. *Princeton University Press*, 1957.
- [7] R.E. Bellman. Adaptive Control Processes : a Guided Tour. *Princeton University Press*, 1961.
- [8] R.E. Bellman and S. Dreyfus. Applied Dynamic Programming. *Princeton University Press*, 1961.
- [9] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-Ph. Vial. Retailer-Supplier Flexible Commitments Contracts : a Robust Optimization Approach. *Manufacturing and Service Operations Management*, 7 :248–271, 2005.

- [10] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming*, 99(2) :351–376, 2004.
- [11] A. Ben-Tal and A. Nemirovski. Selected Topics in Robust Convex Optimization. *Mathematical Programming : Series A and B*, 112 :125–158, 2007.
- [12] D.P. Bertsekas. Convergence of Discretization Procedures in Dynamic Programming. *IEEE transactions on automatic control*, AC-20 :415–419, 1975.
- [13] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena scientific, 1995.
- [14] D.P. Bertsekas and D.A. Castanon. Adaptative Aggregation Methods for Infinite Horizon Dynamic Programming. *IEEE transactions on automatic control*, AC-34 :589–598, 1989.
- [15] D.P. Bertsekas and S.E. Shreve. *Stochastic Optimal Control : the Discrete Time Case*. Academic Press, 1978.
- [16] D. Bertsimas and A. Thiele. A Robust Optimization Approach to Inventory Theory. *Operations Research*, 54(1) :150–168, 2006.
- [17] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag New York, Inc., 1997.
- [18] J.R. Birge. Abridged Nested Decomposition for Multistage Stochastic Linear Programming. Technical report, 1999.
- [19] J. Blomvall and A. Shapiro. Solving Multistage Asset Investment Problems by the Sample Average Approximation Method. Technical report, Georgia Institute of Technology, Atlanta, GA, 2004, revised 2005.
- [20] P. Bonami and M. A. Lejeune. An Exact Solution Approach for Portfolio Optimization Problems under Stochastic and Integer Constraints. Technical report, 2007.
- [21] S. Brignol and A. Renaud. A New Model for Stochastic Optimization of Weekly Generation Schedules. In *In Proc. Fourth International Conf. on Advances in Power System Control, volume 2*, pages 656–661, 1997.
- [22] P. Carpentier. Méthodes Numériques en Optimisation Stochastique, Cours du Master MMME, Filière Optimisation et Commande. Technical report, Université Paris 1, 2005.
- [23] P. Carpentier, G. Cohen, J.C. Culioli, and A. Renaud. Stochastic Optimization of Unit Commitment : a New Decomposition Framework. *IEEE Trans. Power Syst.*, 11 :1067–1073, 1996.
- [24] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost Horizons and Certainty Equivalents : an Approach to Stochastic Programming of Heating Oil. *Management Science*, 4 :235–263, 1958.
- [25] F. Chatelin. *Iterative Aggregation/Desaggregation Methods*, chapter 2.1, pages 199–207. Elsevier Science publishers, Amsterdam, Netherlands, 1984.
- [26] G.B. Dantzig. Linear Programming under Uncertainty. *Management Science*, 1 :197–206, 1955.
- [27] T. Homem de Mello. On Rates of Convergence for Stochastic Optimization Problems under Non-i.i.d. Sampling. Technical report, Dept. of Industrial Engineering and Management Sciences, Northwestern University, 2006.
- [28] E.V. Denardo. *Dynamic Programming Models and Applications*. Prentice-Hall, 1982.

- [29] D. Dentcheva and W. Römisch. Optimal Power Generation under Uncertainty via Stochastic Programming. *Stochastic Programming Methods and Technical Applications*, 458 :22–56, 1998.
- [30] N. Van Dijk. On truncations and Perturbations of Markov Decision Problems with an Application to Queuing Network Overflow Control. *Annals of Operations Research*, 29 :515–536, 1991.
- [31] J.C. Dodu, M. Goursat, A. Hertz, J.-P. Quadrat, and M. Viot. Méthodes de Gradient Stochastique pour l’Optimisation des Investissements dans un Réseau Electrique. In *Série C, Mathématique, Informatique, number 2*, pages 133–167. 1981.
- [32] J. Dupačová, N. Growe-Kuska, and W. Römisch. Scenario Reduction in Stochastic Programming. An Approach using Probability Metrics. *Mathematical Programming*, 95(3) :493–511, 2003.
- [33] M. Dyer and L. Stougie. Computational Complexity of Stochastic Programming Problems. *Mathematical Programming*, 106(3) :423–432, 2006.
- [34] A. Prékopa. *Stochastic Programming*. Kluwer, Dordrecht, Boston, 1995.
- [35] Y. Ermoliev. Stochastic Quasi-Gradient Methods. In Y. Ermoliev and R.J.-B.Wets, editors, *Numerical Techniques for Stochastic Optimization*, pages 143–185. Springer Verlag, Berlin, 1988.
- [36] O. Fiedler and W. Römisch. Stability in Multistage Stochastic Programming. *Annals of Operations Research*, 56 :79–93, 1995.
- [37] B.L. Fox. Finite-State Approximations to Denumerable-State Dynamic Programs. *Journal of Mathematical Analysis and Applications*, 34 :665–670, 1971.
- [38] K. Frauendorfer and J. Güssow. Stochastic Multistage Programming in the Optimization and Management of a Power System. In K. Marti, editor, *Stochastic Optimization Techniques*. Springer Verlag, 2002.
- [39] S. J. Garstka and R. J.-B. Wets. On Decision Rules in Stochastic Programming. *Mathematical Programming*, 7 :117–143, 1974.
- [40] A. Gjelsvik, M.M. Belsnes, and A. Haugstad. An Algorithm for Stochastic Medium-Term Hydrothermal Scheduling under Spot Price Uncertainty. In *In Proceedings 13th Power System Computation Conference*, 1999.
- [41] R. Gollmer, M.P. Nowak, W. Römisch, and R. Schultz. Unit Commitment in Power Generation - A Basic Model and Some Extensions. *Annals of Operations Research*, 96 :167–189, 2000.
- [42] B.G. Gorenstin, N. M. Campodonico, J. P. Costa, and M. V. F. Pereira. Power System Expansion Planning under Uncertainty. *IEEE Trans. Power Syst.*, 8 :129–136, 1993.
- [43] M. Goursat, J.P. Quadrat, and M. Viot. Stochastic Gradient Methods for Optimizing Electrical Transportation Networks. In *Conference on Stochastic Optimization, Kiev*, 1984.
- [44] N. Gröwe-Kuska, H. Heitsch, and W. Römisch. Scenario Reduction and Scenario Tree Construction for Power Management Problems. In A. Borghetti, C.A. Nucci, and M. Pao-lone, editors, *Power Tech Proceedings*. IEEE, Bologna, 2003.

- [45] N. Gröwe-Kuska, K.C. Kiwiel, M.P. Nowak, W. Römisch, and I. Wegner. Power Management in a Hydro-Thermal System under Uncertainty by Lagrangian Relaxation. In A. Borghetti, C.A. Nucci, and M. Paolone, editors, *IMA Volumes in Mathematics and its Applications*, volume 128, pages 39–70. Springer-Verlag, New-York, 2002.
- [46] N. Gröwe-Kuska and W. Römisch. *Stochastic Unit Commitment in Hydro-Thermal Power Production Planning*, chapter 30. MPS-SIAM Series in Optimization, 2005.
- [47] H. Heitsch and W. Römisch. Scenario Reduction Algorithms in Stochastic Programming. *Computational optimization and applications*, 24 :187–206, 2003.
- [48] H. Heitsch and W. Römisch. Generation of Multivariate Scenario Trees to Model Stochasticity in Power Management. *Computational Optimization and Applications*, 24(2-3) :187 – 206, February-March 2003.
- [49] R. Henrion. Introduction to Chance-Constrained Programming. *Tutorial paper for the Stochastic Programming Community Home Page*, 2004.
- [50] R. Henrion. Optimization Problems with Probabilistic Constraints. *10th International Conference on Stochastic Programming, Tucson*, 2004.
- [51] R. Henrion and C. Strugarek. Convexity of Chance Constraints with Independent Random Variables. *In Stochastic Programming E-Print Series*, 9, 2006.
- [52] K. Hinderer. *Foundations of Non-Stationnary Dynamic Programming with Discrete Time Parameter*. Springer-Verlag, 1970.
- [53] C.C. Holt, F. Modigliani, J.F. Muth, and H. Simon. *Planning Production, Inventories, and Work Force*. Prentice-Hall, New Jersey, 1960.
- [54] A. Juditsky, G. Lan, A. Nemirovski, and A. Shapiro. Stochastic Approximation Approach to Stochastic Programming. *SIAM Journal on Optimization*, 2007 (submitted).
- [55] P. Kall. *Stochastic Linear Programming*. Springer-Verlag, Berlin, 1976.
- [56] E. Kiefer and J. Wolfowitz. Stochastic Estimation of the Maximum of a Regression Function. *Ann. Math. Statist.*, 23, 1952.
- [57] M. Koivu. Variance Reduction in Sample Approximations of Stochastic Programs. *Math. Prog.*, 103 :463–485, 2005.
- [58] H.J. Kushner and D.S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer-Verlag, 1976.
- [59] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, 1992.
- [60] J. T. Linderoth, A. Shapiro, and S. J. Wright. The Empirical Behavior of Sampling Methods for Stochastic Programming. *Annals of Operations Research*, 142 :215–241, 2006.
- [61] W.K. Mak, D.P. Morton, and R.K. Wood. Monte-Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs. *Operation Research Letters*, 24 :47–56, 1999.
- [62] L.G.B. Marzano. *Otimização de Portfolio de Contratos de Energia Sistemas Hidrotermicos com Despacho Centralizado*. PhD thesis, Pontificia universidade catolica de Rio de Janeiro, 2004.

- [63] R. Mendelssohn. An Iterative Procedure for Markov Decision Processes. *Operations Research*, 30 :62–73, 1982.
- [64] B.L. Miller and H.M. Wagner. Chance Constrained Programming with Joint Constraints. *Operations Research*, 13 :930–945, 1965.
- [65] G.L. Nemhauser. *Introduction to Dynamic Programming*. Wiley, 1966.
- [66] A. Nemirovski and A. Shapiro. Scenario Approximations of Chance Constraints. Preprint, School of Industrial and Systems Engineering, Georgia Institute of Technology, 2004.
- [67] A. Nemirovski and A. Shapiro. Convex Approximations of Chance Constrained Programs. *SIAM Journal on Optimization*, 17 :969–996, 2006.
- [68] A. Nemirovski and D. Yudin. *Problem Complexity and Method Efficiency in Optimization*. 1983.
- [69] Y. Nesterov and J.-Ph. Vial. Confidence Level Solutions for Stochastic Programming. *Automatica*, 2008 (à paraître).
- [70] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, 1992.
- [71] R. Nürnberg and W. Römisch. A Two-Stage Planning Model for Power Scheduling in a Hydro-Thermal System under Uncertainty. *Optimization and Engineering*, 3 :355–378, 2002.
- [72] T. Numnonda, U.D. Annakkage, and N.C. Pahalawatththa. Unit Commitment using Stochastic Optimization. In *Proc. 1996 International Conference on Intelligent Systems Applications to Power Systems*, 8 :428–433, 1996.
- [73] T. Pennanen and M. Koivu. Epi-Convergent Discretizations of Stochastic Programs via Integration Quadratures. *Numerische Mathematik*, 100 :141–163, 2005.
- [74] M. V. F. Pereira and L. M. V. G. Pinto. Multi-Stage Stochastic Optimization Applied to Energy Planning. *Mathematical Programming*, 52 :359–375, 1991.
- [75] M.V.F. Pereira, N. Campodonico, and R. Kelman. Application of Stochastic Dual Dynamic Programming and Extensions to Hydrothermal Scheduling. Technical report, PSRI, 1999.
- [76] G. Ch. Pflug. Scenario Tree Generation for Multiperiod Financial Optimization by Optimal Discretization. *Mathematical programming*, 89, number 2 :251–271, 2001.
- [77] E.L. Plambeck, B.R. Fu, S.M. Robinson, and M. Suri. Sample-Path Optimization of Convex Stochastic Performance Functions. *Math. Prog.*, B 75 :137–176, 1996.
- [78] B.T. Polyak. Convergence and Convergence Rate of Iterative Stochastic Algorithms. *Automatica i Telemekhanika*, 12 :83–94, 1976.
- [79] B.T. Polyak. Subgradient Methods : a Survey of Soviet Research. *Pergamon Press*, 1978.
- [80] B.T. Polyak and Y.Z. Tsypkin. Pseudogradient Adaptation and Training Algorithms. *Automatica i Telem*, 3, 1973.
- [81] A. Prékopa. On Probabilistic Constrained Programming. In Princeton University Press, editor, *Proceedings of the Princeton Symposium on Mathematical Programming*, pages 113–138, 1970.
- [82] A. Prékopa. Logarithmic Concave Measures with Applications to Stochastic Programming. *Acta Scientifica Mathematica*, 32 :301–316, 1971.

- [83] A. Prékopa. Contributions to the Theory of Stochastic Programming. *Mathematical Programming*, 4 :202–221, 1973.
- [84] M.L. Puterman. *Markov Decision Processes*. John Wiley and sons, 1994.
- [85] W. Römisch. Stability of Stochastic Programming Problems. In *Handbooks in Operations Research and Management, Stochastic Programming*, volume 10, pages 483–554. Elsevier, Amsterdam, 2003.
- [86] W. Römisch and R. Schultz. Stability Analysis for Stochastic Programs. *Annals of Operations Research*, 30 :241–266, 1991.
- [87] W. Römisch and R. Schultz. Stability of Solutions for Stochastic Programs with Complete Recourse. *Mathematics of Operations Research*, 18 :590–609, 1993.
- [88] H. Robbins and S. Monro. A Stochastic Approximation Method. *Ann. Math. Statist.*, 22 :400–407, 1951.
- [89] S.M. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, 1983.
- [90] T.A. Rotting and A. Gjelsvik. Stochastic Dual Dynamic Programming for Seasonal Scheduling in the Norwegian Power System. *IEEE transactions on power systems*, 7 :273–279, 1992.
- [91] A. Ruszczyński. *Decomposition Methods*, volume 10 of *In Handbooks in Operations Research and Management Science*. Elsevier, North-Holland, Amsterdam, 2003.
- [92] A. Ruszczyński and A. Shapiro. *Stochastic Programming*, volume 10 of *In Handbooks in Operations Research and Management Science*. Elsevier, North-Holland, Amsterdam, 2003.
- [93] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A Stochastic Programming Approach for Supply Chain Network Design under Uncertainty. *Eur. J. Oper. Res.*, 167 :96–115, 2005.
- [94] L.I. Sennott. The Computation of Average Optimal Policies in Denumerable State Markov Decision Chains. *Advances in Applied Probability*, 29 :114–137, 1997.
- [95] L.I. Sennott. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley and sons, New-York, NY, 1999.
- [96] A. Shapiro. Complexity of Two and Multi-Stage Stochastic Programming Problems. Conference transparencies, 2005.
- [97] A. Shapiro. On Complexity of Multistage Stochastic Programs. *Operations Research Letters*, 34 :1 – 8, 2006.
- [98] A. Shapiro. Stochastic Programming Approach to Optimization under Uncertainty. Working paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332- Atlanta, Georgia 30332-0205, USA, 2006 (to appear in *Mathematical Programming*).
- [99] A. Shapiro and A. Nemirovski. On Complexity of Stochastic Programming Problems. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA, 2004.
- [100] A. Shapiro and A. Philpott. A Tutorial on Stochastic Programming. Technical report, 2007.

BIBLIOGRAPHIE

- [101] A. Shapiro and A. Ruszczyński. *Lectures on Stochastic Programming*. 2008.
- [102] S. Takriti, C. Supatgiat, and L. Wu. Coordinating Fuel Inventory and Electric Power Generation under Uncertainty. *IEEE Trans. Power Syst.*, 16 :603–608, 2001.
- [103] J. Thénié and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : a Heuristic Approach. *Automatica*, 2008 (to appear).
- [104] P. Varaiya and J.-B. Wets. *Stochastic Dynamic Optimization Approaches and Computation*. Kluwer Academic Publisher, 1989.
- [105] B. Verweij, S. Ahmed, A.J. Kleywegt, G. Nemhauser, and A. Shapiro. The Sample Average Approximation Method Applied to Stochastic Routing Problems : a Computational Study. *Comput. Optim. Appl.*, 24 :289–333, 2003.
- [106] R.J.-B. Wets. Optimization. In G.L. Nemhauser, A.H.G. Rinnoy Kan, and M.J. Todd, editors, *Handbooks in Operations Research and Management, Stochastic Programming*, volume 1, pages 573–629. North-Holland Publishing Co, Amsterdam, 1989.
- [107] D.J. White. On Truncations and Perturbations of Markov Decision Problems with an Application to Queuing Network Overflow Control. *Journal of Mathematical analysis and Applications*, 74 :292–295, 1980.
- [108] P.J. Wong. An Approach to Reducing the Computing Time for Dynamic Programming. *Operations Research*, 18 :181–185, 1970.

Deuxième partie

Etude et application de l'optimisation robuste

Chapitre 3

Introduction à l'optimisation robuste linéaire

Sommaire

3.1	Introduction	76
3.2	Problème d'optimisation linéaire avec coefficients incertains	78
3.3	Satisfaction en probabilité d'une contrainte d'inégalité	79
3.3.1	Distribution normale	80
3.3.2	Distribution uniforme	80
3.3.3	Distribution générale	82
3.4	Solution robuste pour un ensemble d'incertitude	83
3.4.1	Contrainte robuste	84
3.4.2	Equivalent robuste d'une contrainte d'inégalité	86
3.5	Probabilité de satisfaction de la contrainte pour une solution robuste	90
3.5.1	Ensemble d'incertitude $\Xi = B(0, k)_2$	91
3.5.2	Ensemble d'incertitude $\Xi = B(0, 1)_\infty \cap B(0, k)_1$	93
3.6	Éléments pour définir l'ensemble d'incertitude	94
3.6.1	Une approche heuristique basée sur un aléa gaussien	94
3.6.2	Equivalent robuste	95
3.7	Problème général d'optimisation linéaire	97
3.7.1	Equivalent robuste d'une contrainte d'égalité	97
3.7.2	Performance moyenne et performance garantie	98
3.8	Problèmes dynamiques	99
3.8.1	Problèmes avec recours fixe	101
3.8.2	Règle de décision linéaire par morceaux avec coefficients de recours fixes	101
3.8.3	Problèmes avec coefficient de recours incertain	102
3.9	Extension du concept de robustesse : violation contrôlée des contraintes	103
3.9.1	Définition formelle de la robustesse étendue	103
3.9.2	Robustesse étendue dans le cadre de la programmation linéaire	104
3.10	Conclusion	107
	Bibliographie	108

Ce chapitre est une introduction à l'optimisation robuste, limitée au cadre linéaire. En effet, les grands thèmes de l'optimisation robuste sont basés sur la dualité. Or, la dualité dans le cadre de la programmation linéaire a l'extrême avantage d'avoir des écritures moins lourdes que dans le cadre de la programmation conique. Bien que restreint, nous avons donc fait le choix délibéré de développer cette introduction dans le cadre linéaire afin d'être le plus clair possible. Ce chapitre reprend des principes de base (liens avec les contraintes en probabilité, obtention d'équivalents robustes) mais détaille également des extensions possibles (cas dynamique via des règles de décision, robustesse étendue,...). De plus, ce chapitre contient un énoncé (et sa preuve) de la proposition 3.1, permettant de borner la probabilité de non-satisfaction de la contrainte, plus fort que ce que la littérature fournit actuellement. Cette introduction, permettant de mettre en place des concepts utilisés dans la suite de cette thèse, a donné lieu à une note interne EDF [40] et a été soumis à [41].

3.1 Introduction

L'*optimisation robuste* traite de problèmes d'optimisation dans lesquels les coefficients des contraintes et de l'objectif sont incertains. Son but est de fournir une solution opérationnelle à des problèmes de cette classe, en des temps de calcul raisonnables, là où les approches traditionnelles, basées sur un modèle probabiliste, se heurtent à des difficultés numériques considérables. En effet, il est en général très difficile, voire impossible, de formuler un modèle probabiliste pertinent pour les paramètres incertains du problème. Lorsque cela est malgré tout possible, le calcul d'espérances mathématiques avec des distributions multidimensionnelles peut être numériquement très lourd, voire impossible à effectuer avec une précision raisonnable. Et si ces difficultés sont surmontées, le problème d'optimisation est alors d'une telle dimension ou d'une complexité telle qu'il est impossible à résoudre, en particulier lorsqu'il s'agit de problèmes dynamiques avec plus de 2 périodes.

L'optimisation robuste contourne ces difficultés en stipulant au départ des hypothèses minimales et non probabilistes sur le domaine de variation des paramètres incertains. Le calcul d'espérances mathématiques est ainsi évacué; il est remplacé par la notion de «pire des cas». Cela conduit à des problèmes d'optimisation de même nature que la version déterministe du problème original et de taille modérément supérieure. L'optimisation robuste est donc opérationnelle, là où les approches alternatives se heurtent à des difficultés souvent insolubles de formulation et de résolution numérique, en particulier lorsqu'il s'agit de problèmes multi-étapes et/ou avec aléas multi-dimensionnels.

L'optimisation robuste propose donc une approche de protection contre le «pire des cas». Si le pire des cas est choisi parmi un ensemble trop vaste de possibles, l'exigence peut être trop sévère pour qu'il existe une solution admissible. L'idée est donc de ne choisir le pire cas que dans un ensemble restreint de possibles appelé *ensemble d'incertitude*. Cela exclut des cas extrêmes qu'une analyse probabiliste aurait sans doute déclaré hautement improbables. Cette formulation par le pire des cas conduit à remplacer la satisfaction d'une contrainte déterministe, par la satisfaction d'un nombre infini de contraintes, chacune de ces contraintes correspondant à une réalisation particulière de l'aléa dans l'ensemble d'incertitude. Le problème d'optimisation qui en

résulte appartient à la programmation mathématique semi-infinie dont la complexité semble bien éloignée de l'objectif de simplicité et de facilité de calculs évoquée plus haut. Ce n'est pourtant pas le cas, car sous des hypothèses appropriées sur les ensembles d'incertitude (par exemple, des polyèdres ou des ellipsoïdes) et sur les contraintes initiales (par exemple la linéarité), il s'avère possible de remplacer cet ensemble infini de contraintes par un nombre fini d'autres contraintes. L'outil qui permet cette transformation est un théorème de dualité forte. Nous obtenons grâce à lui une formulation explicite de l'*équivalent robuste* comme un problème de programmation convexe de dimension raisonnable que nous pouvons résoudre par des algorithmes efficaces.

La première formulation d'une solution robuste à un problème de programmation mathématique avec coefficients incertains peut être attribuée à Soyster [38] au début des années 70. Le domaine a pris son essor beaucoup plus tard, dans la fin des années 90, grâce au travail de pionniers de Ben-Tal et Nemirovski [7], mais également de El-Gahoui et Lebret [25]. Depuis, l'optimisation robuste est devenue un des domaines de recherche en optimisation des plus actifs. Les principaux développements théoriques se trouvent dans les articles de Ben-Tal et co-auteurs [8, 6, 4, 11] ou Bertsimas [13, 12].

L'approche par optimisation robuste semble orthogonale à une vision classique basée sur le calcul des probabilités. Néanmoins, il existe une parenté évidente entre la satisfaction d'une contrainte pour le pire des cas choisi dans un ensemble d'incertitude et l'optimisation sous contrainte en probabilité [21, 22, 19] connue sous le nom de *Chance-constrained programming*. Nous nous servons de cette proximité pour introduire le concept d'ensemble d'incertitude et le rendre plus intuitif.

L'objectif de ce chapitre est de donner une présentation simple des principaux concepts de l'optimisation robuste et de détailler leur mise en œuvre. Pour cette raison, nous avons choisi le cadre de la programmation linéaire et de la modélisation des ensembles d'incertitude par des polyèdres. En faisant l'hypothèse que la formulation déterministe initiale du problème est un programme linéaire, nous détaillons les transformations à effectuer pour obtenir une *contre-partie robuste* formulée elle-aussi comme un programme linéaire. Cette contre-partie robuste introduit des variables et des contraintes auxiliaires, mais en nombre suffisamment réduit pour que le problème transformé soit traitable par les solveurs actuellement disponibles sur le marché, que ce soient des logiciels libres ou des produits commerciaux. Même sous cette forme, l'optimisation robuste linéaire apporte une réponse concrète, opérationnelle et convaincante à un très grand nombre de problèmes de décision dans l'incertain. Le nombre d'applications et la variété des domaines qu'elle couvre ne cessent de croître. Nous pouvons citer par exemple les problèmes de finance et de mesures de risque [15, 18, 20, 27], de gestion des stocks et de chaînes d'approvisionnement [1, 5, 14], de télécommunications [34] ou de gestion de la production d'électricité d'une vallée hydraulique [2]. Cette approche peut également s'appliquer en complément de techniques d'optimisation comme les contraintes en probabilités [19], la programmation dynamique [30], ou la programmation stochastique [23].

Après avoir défini le problème d'optimisation linéaire avec coefficients incertains, nous motivons, dans la section 3.3, l'approche robuste à partir du concept de satisfaction en probabilité d'une contrainte. La section suivante 3.4 définit formellement une solution robuste. Elle est prolongée par la section 3.5 qui montre que sous des hypothèses probabilistes minimales, il est parfois possible de borner la probabilité de non satisfaction d'une contrainte par la solution robuste.

Ces résultats confirment la proximité de l'approche robuste avec l'approche de satisfaction des contraintes en probabilité. La section 3.6 propose une approche heuristique à la détermination de l'ensemble d'incertitude, tandis que la section 3.7 traite du cas des contraintes d'égalité et de la possibilité de prendre en compte un critère de coût moyen en même temps qu'une borne sur la pire des performances. La section 3.8 aborde le cas très important des problèmes dynamiques. Nous y introduisons le concept de règle de décision linéaire que nous étendons à celui de règle de décision linéaire par morceaux. La section 3.9 discute une extension récente du concept de robustesse, qui permet de contrôler le comportement d'une contrainte pour des réalisations hors du domaine d'incertitude. Enfin, la section 3.10 propose une brève conclusion.

3.2 Problème d'optimisation linéaire avec coefficients incertains

Le problème canonique qui nous concerne a la forme suivante :

$$\max_x \quad \sum_{j=1}^n c_j x_j \quad (3.1a)$$

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m_1 \quad (3.1b)$$

$$\sum_{j=1}^n g_{ij} x_j = h_i, \quad i = m_1 + 1, \dots, m_2. \quad (3.1c)$$

Ce problème simple se complique énormément si nous nous trouvons dans une situation où les coefficients de la fonction objectif et des contraintes sont aléatoires. Que signifie alors une contrainte d'inégalité ? Une contrainte d'égalité ? Faut-il, pour une solution donnée x , que les contraintes soient satisfaites presque toujours ? Comment évaluer les performances de la fonction économique ? La programmation robuste propose une approche qui s'apparente à un critère de maximisation du profit minimum garanti, critère dit du « maximin », à ceci près que le maximin ne s'applique pas sur la totalité des événements possibles mais seulement sur une fraction choisie d'entre eux.

Dans cette approche, la fonction objectif et les contraintes d'égalité ont un traitement à part. Pour la fonction objectif, nous procédons à une transformation simple qui permet de la traiter comme une contrainte d'inégalité. Nous introduisons pour cela une variable auxiliaire z , représentant un minorant sur la valeur de l'objectif à maximiser, et la contrainte d'inégalité

$$-\sum_{j=1}^n c_j x_j + z \leq 0.$$

Cela donne le problème équivalent

$$\begin{aligned} \max_{x,z} \quad & z \\ & - \sum_{j=1}^n c_j x_j + z \leq 0 \\ & \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m_1 \\ & \sum_{j=1}^n g_{ij} x_j = h_i, \quad i = m_1 + 1, \dots, m_2. \end{aligned}$$

3.3 Satisfaction en probabilité d'une contrainte d'inégalité

Dans sa définition et dans sa mise en œuvre, l'optimisation robuste ne fait pas appel au calcul des probabilités. Pourtant nous choisissons de motiver le concept d'optimisation robuste par une approche particulière de la programmation stochastique due à Charnes, Cooper et Symonds [22, 21], dénommée par ces auteurs *Chance-Constrained Programming*, ce qui peut se traduire en français par *programmation par satisfaction des contraintes en probabilité*. Cette approche de la programmation stochastique fait l'objet d'une abondante littérature (voir par exemple [36]). Notre objectif n'est pas d'en faire un exposé, mais seulement d'utiliser le concept intuitif d'une probabilité de satisfaction d'une contrainte pour introduire l'optimisation robuste et mettre l'accent sur les similitudes et les différences entre les deux approches.

Dans l'approche par programmation par satisfaction des contraintes en probabilité, chaque contrainte du problème d'optimisation est traitée séparément. Sans perte de généralité, nous considérons une contrainte du type

$$\sum_{j=1}^n a_j(\tilde{\xi}) x_j \leq b$$

où l'incertitude ne porte que sur le membre de gauche. (Nous pouvons nous ramener à ce cas lorsque b est aussi incertain, en introduisant une variable auxiliaire $x_{n+1} \equiv 1$ et en écrivant $b(\xi) = -a_{n+1}$. Nous obtenons alors la contrainte $\sum_{j=1}^n a_j(\tilde{\xi}) x_j + a_{n+1}(\tilde{\xi}) x_{n+1} \leq 0$.) Dans l'expression de la contrainte, $\tilde{\xi}$ est une variable aléatoire avec une distribution connue et x est la variable de décision. Pour la simplicité de l'exposé, nous nous contentons de dire que les réalisations ξ de $\tilde{\xi}$ prennent leurs valeurs dans un espace $\Omega \in \mathbb{R}^p$ et que cet espace est probabilisé.

Nous dirons que la contrainte est satisfaite en probabilité, avec une probabilité $1 - \alpha$, si x vérifie la condition

$$\text{Prob}(\tilde{\xi} \mid \sum_{j=1}^n a_j(\tilde{\xi}) x_j \leq b) \geq 1 - \alpha. \quad (3.2)$$

La simplicité de la formule cache de grandes difficultés. Le domaine des solutions x réalisables pour (3.2) n'est en général pas convexe, voire non connexe. De plus, sa définition est implicite

et sa représentation par des fonctions non-linéaires directement gérables par programmation mathématique n'est pas évidente.

Pour illustrer le propos, considérons l'inégalité simple

$$a_1x_1 + a_2x_2 \leq b$$

où les paramètres a_1 et a_2 sont incertains. Nous supposons connue la distribution de probabilité de ces paramètres. Nous nous proposons de traiter deux cas simples, celui d'une distribution normale et celui d'une distribution uniforme, avant de discuter le cas général.

3.3.1 Distribution normale

Commençons par le cas d'une loi normale en écrivant

$$a_i(\tilde{\xi}) = \mu_i + \sigma_i \tilde{\xi}_i, \quad i = 1, 2$$

où $\tilde{\xi}_1$ et $\tilde{\xi}_2$ représentent deux variables aléatoires indépendantes suivant chacune une loi normale centrée réduite. Les paramètres μ_i et σ_i sont les deux premiers moments de la distribution des $a_i(\tilde{\xi})$. Or, la loi normale est une loi stable : ainsi, la combinaison linéaire $a_1(\tilde{\xi}_1)x_1 + a_2(\tilde{\xi}_2)x_2$ suit une loi normale de moyenne $\mu = \mu_1x_1 + \mu_2x_2$ et d'écart-type $\sigma = \sqrt{\sigma_1^2x_1^2 + \sigma_2^2x_2^2}$. Soit $k_\alpha = \Phi^{-1}(1 - \alpha)$, où Φ est la loi de distribution de la loi normale centrée réduite. La condition

$$\mu_1x_1 + \mu_2x_2 + k_\alpha \sqrt{\sigma_1^2x_1^2 + \sigma_2^2x_2^2} \leq b \quad (3.3)$$

est l'équivalent algébrique de (3.2). Dans ce cas, nous avons obtenu, contrairement à ce que nous pouvions craindre, une expression explicite de la contrainte en probabilité ; de plus, le membre de gauche de (3.3) est une fonction convexe. En terme de complexité, le prix à payer lorsque nous passons de la contrainte déterministe à la contrainte en probabilité n'est pas nul, puisque nous substituons à une contrainte linéaire une contrainte non-linéaire, mais il est faible. En effet, la contrainte peut se récrire :

$$\mu_1x_1 + \mu_2x_2 + k_\alpha t \leq b$$

avec

$$\sqrt{\sigma_1^2x_1^2 + \sigma_2^2x_2^2} \leq t.$$

Cette dernière contrainte décrit un cône quadratique convexe. Les problèmes d'optimisation comportant des contraintes linéaires et des contraintes coniques quadratiques sont « traitables », c'est-à-dire qu'ils peuvent être résolus avec une complexité polynomiale par des méthodes de points intérieurs [33].

3.3.2 Distribution uniforme

Considérons maintenant le cas apparemment plus simple où les paramètres suivent chacun une loi uniforme sur l'intervalle $[a_i^n - a_i^v, a_i^n + a_i^v]$, $i = 1, 2$. Les paramètres a_i^n sont les valeurs nominales alors que les a_i^v donnent l'amplitude des variations possibles. Pour l'exemple numérique nous

3.3. SATISFACTION EN PROBABILITÉ D'UNE CONTRAINTE D'INÉGALITÉ

choisissons $b = 1$, $a_i^n = 1$ et $a_i^v = 1.5$, $i = 1, 2$. Si $\tilde{\xi}_1$ et $\tilde{\xi}_2$ sont deux variables aléatoires indépendantes de loi uniforme U sur $[-1, 1]$, nous pouvons écrire la variable aléatoire $a(\tilde{\xi})$ sous la forme

$$a_i(\tilde{\xi}) = a_i^n + a_i^v \tilde{\xi}_i, \quad i = 1, 2.$$

Nous obtenons la contrainte en probabilité

$$\text{Prob} \left(\tilde{\xi} \mid \sum_{i=1}^2 (a_i^v x_i) \tilde{\xi}_i \leq b - \sum_{i=1}^2 a_i^n x_i \right) \geq 1 - \alpha.$$

La probabilité peut se calculer par une intégrale double sur un carré. C'est chose facile pour un x donné, mais il faut distinguer plusieurs cas, ce qui ne permet pas de caractériser le domaine défini par la contrainte en probabilité comme une contrainte fonctionnelle classique $g(x) \leq 0$, où $g(x)$ serait une fonction donnée sous forme explicite.

Il existe tout de même deux valeurs de probabilité pour lesquelles la réponse est immédiate. Le premier cas est celui où le point x sature la contrainte déterministe : $a_1^n x_1 + a_2^n x_2 = b$. Dans ce cas la probabilité cherchée est donnée par

$$\text{Prob}(\tilde{\xi} \mid (a_1^v x_1) \tilde{\xi}_1 + (a_2^v x_2) \tilde{\xi}_2 \leq 0) \geq 1 - \alpha.$$

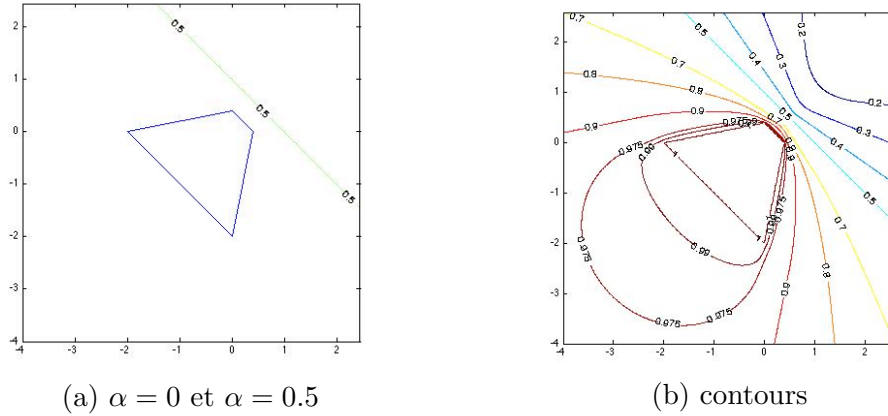
Quels que soient x_1 et x_2 (pour autant que $a_1^v x_1$ et $a_2^v x_2$ ne soient pas simultanément nuls), l'ensemble des ξ réalisant l'événement est l'intersection d'un demi-espace passant par l'origine et du carré $[-1, 1] \times [-1, 1]$. La surface ainsi construite est exactement la moitié de la surface du carré. La probabilité est donc $1/2$. Nous pouvons facilement montrer que les seuls points x réalisant une probabilité $1/2$ sont tous sur la droite correspondant à la contrainte déterministe.

L'autre cas particulier est celui où nous excluons toute violation de la contrainte, soit une probabilité 1 de satisfaction de la contrainte. Les x recherchés doivent demeurer réalisables contre le pire des cas. Pour une inégalité linéaire, le pire cas s'obtiendra quand les variables ξ_i prennent une des deux valeurs extrêmes -1 ou 1 suivant le signe de leur coefficient $a_i^v x_i$. Nous pouvons donc définir l'ensemble recherché par la condition

$$\sum_{i=1}^2 (a_i^n x_i + |a_i^v x_i|) \leq b.$$

Il est facile de voir que cet ensemble est un polyèdre simple, délimité par 4 droites selon le signe de x_1 et x_2 . La figure 3.1(a) illustre les deux ensembles mis en évidence. La figure 3.1(b) donne les contours des ensembles admissibles définis par la contrainte en probabilité pour des valeurs intermédiaires de α .

Nous voyons que malgré la simplicité de l'exemple, l'ensemble réalisable défini par la contrainte en probabilité est complexe. Les courbes de niveau passent par déformation progressive d'un polyèdre, à un ensemble à frontière lisse borné, puis non borné. Pour $\alpha > 0.5$, l'ensemble n'est plus convexe. Dans tous les cas, l'ensemble des points réalisables peut être difficile à appréhender numériquement.


 FIG. 3.1 – Points admissibles pour la contrainte en probabilité avec $1 - \alpha$

3.3.3 Distribution générale

Les difficultés que nous avons rencontrées dans le cas de la distribution uniforme (difficulté à vérifier si la contrainte en probabilité est satisfaite en un x donné et non-convexité de l'ensemble des solutions admissibles) se retrouvent aggravées dans le cas général d'une distribution quelconque. Nous avons vu une exception, celle associée à une loi normale multivariée. La deuxième concerne le cas où la condition $(x, \xi) \in Q$, où Q est un ensemble convexe fermé et où la distribution du vecteur aléatoire $\xi \in \mathbb{R}^p$ est logarithmiquement quasi-concave (cf. [24]). Ce sont d'après [37] les deux seuls cas génériques donnant lieu à des expressions explicites et traitables de l'ensemble de réalisabilité pour la contrainte en probabilité. Henrion dans [28] a toutefois étendu ce résultat au cas des distributions elliptiques symétriques (qui incluent notamment la loi exponentielle, la loi de Student,...).

Dans ces conditions, il est naturel de chercher une approximation de la contrainte en probabilité qui soit traitable par un algorithme de programmation convexe. Parmi les propositions faites dans ce sens, citons en une récente [16] que nous pouvons brièvement illustrer sur notre exemple de contrainte linéaire. Nous supposons comme précédemment que

$$a_i(\tilde{\xi}_i) = \mu_i + \sigma_i \tilde{\xi}_i, \quad i = 1, 2,$$

et que les $\tilde{\xi}_i$ sont des variables aléatoires indépendantes, centrées et réduites. Les deux premiers moments de la variable $z(x) = a_1(\tilde{\xi}_1)x_1 + a_2(\tilde{\xi}_2)x_2$ sont, comme dans l'inégalité (3.3), $\mu(x) = \mu_1 x_1 + \mu_2 x_2$ et $\sigma^2(x) = \sigma_1^2 x_1^2 + \sigma_2^2 x_2^2$. Soit

$$Z_x = \frac{[a_1(\tilde{\xi}_1)x_1 + a_2(\tilde{\xi}_2)x_2] - [\mu_1 x_1 + \mu_2 x_2]}{\sigma_1^2 x_1^2 + \sigma_2^2 x_2^2}.$$

C'est une variable centrée réduite dont la distribution

$$F_x(t) = \text{Prob}(Z_x \leq t)$$

dépend de x . Supposons F_x continue et définissons $t_\alpha(x) = F_x^{-1}(1 - \alpha)$. La contrainte (3.3) peut maintenant s'écrire comme

$$\mu_1 x_1 + \mu_2 x_2 + t_\alpha(x) \sqrt{\sigma_1^2 x_1^2 + \sigma_2^2 x_2^2} \leq b. \quad (3.4)$$

La dépendance de t_α vis-à-vis de x fait que cette formulation n'est pas opérationnelle. Cependant, sous l'hypothèse assez faible que la distribution de $\sum_i a_i(\tilde{\xi}_i)$ est symétrique et que $\alpha > 0.5$, Bonami et Lejeune [16] montrent que (3.4) est bien une contrainte conique quadratique. Ces auteurs donnent de plus des conditions sous lesquelles $t_\alpha(x)$ peut être borné par une constante dépendant de α seulement (voir [37] pour une discussion plus complète). D'autres approximations ont été proposées. Nemirovski et Shapiro [32] utilisent l'approximation de Bernstein introduite par [35] dans le cadre des contraintes en probabilité. Ils obtiennent ainsi une estimation de l'erreur d'approximation.

Une autre possibilité consiste à approximer la contrainte en probabilité du type de (3.2) par un ensemble de N contraintes $\sum_{j=1}^n a_j(\xi^i)x_j \leq b$, $\forall i = 1, \dots, N$, construites à l'aide de N scénarios indépendants ξ^i discrétisant l'incertitude. Moyennant certaines hypothèses sur les distributions, il est possible de garantir, avec un niveau de confiance donné, que la contrainte en probabilité est satisfaite. Pour que cette approche soit numériquement traitable, il faut d'une part pouvoir borner la complexité du problème d'optimisation sur les scénarios générés (en particulier la dépendance de la complexité vis-à-vis du nombre de scénarios) et d'autre part s'assurer que le nombre de scénarios requis pour garantir un certain niveau d'optimalité ne soit pas trop grand. Des travaux récents [17, 31] discutent ces approches en profondeur.

3.4 Solution robuste pour un ensemble d'incertitude

Dans l'approche par contrainte en probabilité, le concept premier est le modèle probabiliste de l'incertitude. Ce modèle est ensuite utilisé pour caractériser l'admissibilité des solutions envisagées. La troisième et dernière étape est l'incorporation de ces données dans un problème d'optimisation. La section précédente donne une idée des difficultés liées à cette approche et des objections que nous pouvons formuler à son encontre. La première objection concerne la pertinence même d'un modèle probabiliste dans un problème d'optimisation impliquant de très nombreux paramètres incertains. Le modèle probabiliste est alors trop complexe pour être utilisable, ou trop arbitraire. La deuxième objection est celle que nous avons mise en évidence dans la section précédente : la caractérisation analytique de l'ensemble admissible pour une contrainte en probabilité est en règle générale impossible et, de plus, les calculs pour vérifier la validité de la contrainte en un point donné peut s'avérer très difficile numériquement. Enfin l'ensemble admissible peut s'avérer non convexe. Nous devons donc nous attendre à devoir résoudre des problèmes d'optimisation non-linéaire généraux de très grande taille pour lesquels il n'existe pas de méthode avec une complexité satisfaisante, comme c'est le cas pour la programmation linéaire conique.

L'approche robuste consiste à renverser l'ordre des priorités. L'optimisation robuste s'appuie sur deux principes directeurs. Le premier est que la formalisation d'un problème d'optimisation dans l'incertain doit conduire à un problème d'optimisation numériquement «traitable», c'est-à-dire à un problème qui peut se résoudre par des méthodes polynomiales de la programmation convexe. Le deuxième principe est qu'il faut substituer au modèle probabiliste, lui-même incertain et souvent complexe, le concept d'ensemble d'incertitude, c'est-à-dire un ensemble de valeurs des paramètres incertains envers lesquels toute solution du problème d'optimisation doit préserver la réalisabilité des contraintes. Cet ensemble sera choisi dans une classe restreinte : ensembles

polyédriques, ou ensembles décrits par des contraintes coniques quadratiques. Nous verrons comment cette approche conduit effectivement à des problèmes «traitables». La dernière étape consiste à tenter de réconcilier l'approche robuste et l'approche par contrainte en probabilité en estimant *a posteriori* la probabilité que la contrainte soit satisfaite pour la solution robuste trouvée. Les deux points de vue semblent alors se rejoindre. Malgré tout, il reste une différence fondamentale dans cette dernière étape. En effet, la probabilité de satisfaction de la contrainte calculée *a posteriori* n'est en général pas évaluée pour une distribution particulière, mais pour une classe aussi large que possible de distributions.

3.4.1 Contrainte robuste

Explicitons à présent en partie la démarche robuste sur l'exemple introductif de la section précédente. Comme il a été dit, le problème est pris à l'envers. Plutôt que d'assigner des probabilités aux coefficients $a_i(\tilde{\xi}_i)$, nous nous proposons de restreindre arbitrairement le vecteur x à des ensembles simples comme celui associé à la contrainte en probabilité avec $\alpha = 0$. Reprenons le petit exemple numérique de la section précédente. Nous pouvons définir un ensemble par la condition

$$\sum_{i=1}^2 (a_i^n x_i + k|a_i^v x_i|) \leq b, \quad (3.5)$$

où $0 \leq k \leq 1$ est un paramètre arbitraire. Pour $k = 1$ nous retrouvons l'ensemble associé à la contrainte en probabilité avec $\alpha = 0$, alors qu'avec $k = 0$ nous obtenons le demi-espace associé à la contrainte en probabilité avec $\alpha = 0.5$. Les ensembles définis par quelques valeurs intermédiaires de k sont représentés sur une même figure (figure 3.2), en même temps que des ensembles associés avec des contraintes en probabilité. Lorsque k diminue, le domaine réalisable s'agran-

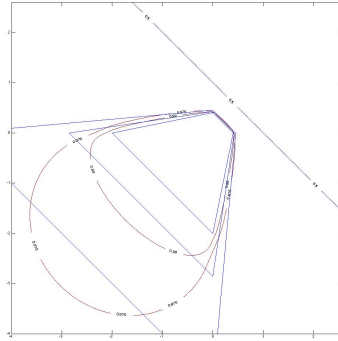


FIG. 3.2 – Contrainte par ensemble d'incertitude

dit. Le polyèdre intermédiaire sur la figure 3.2 correspond à $k = 0.9$; il approxime l'ensemble réalisable associé à la contrainte en probabilité 0.99. Nous disons que les solutions réalisables pour la contrainte par ensemble d'incertitude de niveau k sont des solutions robustes pour le problème initial avec un niveau d'immunisation k .

En remplaçant la contrainte en probabilité avec $\alpha = 0.01$ par la contrainte par ensemble d'incertitude avec $k = 0.9$, nous avons substitué à un ensemble réalisable non-linéaire complexe,

un polyèdre simple. Dans cette nouvelle approche, la contrainte linéaire initiale avec coefficients incertains est remplacée par plusieurs contraintes linéaires. Si la contrainte déterministe était initialement associée à un problème de programmation linéaire, la nouvelle formulation conduit encore à un problème de programmation linéaire. Nous restons donc dans le domaine des problèmes *traitables*, c'est-à-dire des problèmes que nous pouvons résoudre *en temps polynomial* (dans le cas présent, par programmation linéaire), ce qui procure un avantage certain.

Dans le reste de la section, nous allons formaliser l'approche par contrainte d'incertitude, sans nous soucier de lui donner un équivalent traitable comme dans la contrainte (3.5). Pour traiter le cas général où l'incertitude porte sur a et b dans la contrainte $\sum_{i=1}^n a_i x_i \leq b$, nous introduisons la variable de décision auxiliaire $x_{n+1} \equiv 1$ et le paramètre $a_{n+1}(\xi) = -b(\xi)$. Nous récrivons alors la contrainte comme

$$\sum_{i=1}^{n+1} a_i(\xi) x_i \leq 0.$$

A la place de (3.2), nous proposons une formulation plus restrictive, mais plus directe, sur les choix possibles de x . Soit $\Xi \subset \Omega$ un *ensemble dit d'incertitude*. Nous imposons alors que la contrainte soit satisfaite pour toute réalisation dans l'ensemble Ξ :

$$\sum_{i=1}^{n+1} a_i(\xi) x_i \leq 0, \forall \xi \in \Xi. \quad (3.6)$$

Une solution qui satisfait (3.6) est dite *robuste*. Remarquons que si $\text{Prob}(\tilde{\xi} \in \Xi) \geq 1 - \alpha$, la contrainte (3.6) satisfait (3.2). L'inverse n'est pas vrai en général. Est-ce que la nouvelle contrainte est pour autant plus facile à manipuler ? A priori ce n'est pas évident. L'ensemble Ξ peut comporter une infinité de valeurs et la contrainte (3.6) représente en fait une infinité de contraintes. Nous disons que nous sommes dans un cadre de programmation semi-infinie.

Pour rendre la formulation traitable par la programmation mathématique finie, il faudrait pouvoir remplacer la contrainte (3.6) par une contrainte sur les x qui ne fasse pas intervenir la variable aléatoire ξ . Pour cela nous pouvons définir un test de robustesse pour une solution particulière x en résolvant le problème d'optimisation

$$\max_{\xi} \left\{ \sum_{i=1}^{n+1} a_i(\xi) x_i \mid \xi \in \Xi \right\}.$$

Si ce maximum est négatif, la contrainte (3.6) est satisfaite ; la solution x est donc robuste. Soit

$$\xi(x) = \arg \max_{\xi} \left\{ \sum_{i=1}^{n+1} a_i(\xi) x_i \mid \xi \in \Xi \right\}$$

une solution optimale du problème précédent. Nous pouvons utiliser cette fonction pour écrire une contrainte équivalente à (3.6) :

$$\sum_{i=1}^{n+1} a_i(\xi(x)) x_i \leq 0. \quad (3.7)$$

Cette contrainte est appelée *l'équivalent robuste* de (3.6). Elle implique une fonction $\xi(x)$ définie implicitement. En règle générale, une contrainte impliquant une fonctionnelle comme $\xi(x)$ n'est malheureusement pas traitable par les techniques standard de la programmation mathématique. Il semble qu'en passant à la formulation (3.6) nous n'ayons fait que déplacer le problème. Les sections qui suivent vont montrer que sous certaines hypothèses, la théorie de la dualité permet de remplacer la contrainte unique (3.6) faisant intervenir une fonction implicite par un ensemble de contraintes simples directement traitables par la programmation mathématique convexe (linéaire dans le cas de cette note).

3.4.2 Equivalent robuste d'une contrainte d'inégalité

Nous utilisons dans cette section la même convention que dans la section précédente, à savoir que la contrainte est homogène $\sum_{i=1}^{n+1} a_i(\xi)x_i \leq 0$, ce qui sous-entend $x_{n+1} \equiv 1$.

Il est possible d'introduire des hypothèses sur l'aléa qui permettent de remplacer la contrainte implicite (3.6) par un ensemble fini d'inégalités linéaires. Ces hypothèses portent sur les fonctions aléatoires $a_i(\xi)$ et sur l'ensemble d'incertitude Ξ .

Hypothèse 3.1 *L'aléa ξ est une variable aléatoire à valeurs dans \mathbb{R}^p . Les coefficients de la contrainte sont des fonctions affines de l'aléa :*

$$a(\xi) = \bar{a} + P\xi,$$

avec $\bar{a} \in \mathbb{R}^n$ et P est une matrice $n \times p$.

Hypothèse 3.2 *L'ensemble d'incertitude $\Xi \subset \mathbb{R}^p$ a l'une des deux formes suivantes :*

1. *Une boule dans \mathbb{R}^p de la forme*

$$\|\xi\|_r \leq k,$$

avec $r = 1, 2$ ou ∞ .

2. *Un polyèdre $Q\xi \leq t$, où Q est une matrice $m \times p$ et t est un vecteur dans \mathbb{R}^m .*

Les hypothèses 3.1 et 3.2 peuvent s'interpréter comme suit. La variable aléatoire ξ est un sous-jacent¹. L'hypothèse 3.1 exprime que la dépendance entre facteurs aléatoires et le sous-jacent est linéaire. L'hypothèse 3.2 sur l'ensemble d'incertitude n'implique pas grand chose quant à la distribution de ce sous-jacent. Parmi les hypothèses que nous pouvons faire sur le sous-jacent, nous pouvons retenir celle d'un bruit blanc. Nous pourrions alors modéliser la distribution de ξ comme celle d'une loi normale multivariée centrée et réduite (hypothèse de bruit gaussien). Dans ce cas, l'ensemble d'incertitude le plus naturel est la boule dans ℓ_2 . Nous reprendrons cette discussion un peu plus loin.

Il est facile de traiter le cas d'un ensemble d'incertitude décrit par une boule. Nous nous rappelons que pour vérifier si une solution x est robuste, il faut résoudre le problème de maximisation

$$\max_{\xi} \{(\bar{a} + P\xi)^T x \mid \|\xi\|_r \leq k\}.$$

¹Par sous-jacent, nous entendons l'ensemble des facteurs aléatoires qui déterminent la valeur observée des coefficients $a(\xi)$.

Dans ce problème, x est un paramètre fixe. C'est la valeur de la variable x dont nous voulons tester la robustesse. Il suffit de calculer la solution optimale du problème

$$\max_{\xi} \{(P^T x)^T \xi \mid \|\xi\|_r \leq k\}. \quad (3.8)$$

C'est le problème de Cauchy de maximisation d'une forme linéaire sur une boule.

Lemme 3.1

$$\max_{\xi} \{\alpha^T \xi \mid \|\xi\|_r \leq k\} = k \|\alpha\|_s,$$

où $\|\cdot\|_s$ est la norme conjuguée avec $1/r + 1/s = 1$.

Preuve: Nous prouvons le lemme pour $r > 1$. Puisque l'objectif est linéaire, la contrainte $\|u\|_p \leq k$ doit être saturée à l'optimum. Soit $\sigma \geq 0$ le multiplicateur associé à cette contrainte. Les conditions nécessaires d'optimalité du premier ordre s'écrivent

$$\begin{aligned} \alpha_j &= \sigma \left(\sum_{j=1}^p |\xi_j|^r \right)^{\frac{1}{r}-1} |\xi_j|^{r-1} \text{signe}(\xi_j) \\ \left(\sum_{j=1}^p |\xi_j|^r \right)^{\frac{1}{r}} &= k, \end{aligned}$$

avec $\text{signe}(x) = x/|x|$, pour $x \neq 0$.

La première équation impose $\text{signe}(\xi_j) = \text{signe}(\alpha_j)$. La deuxième relation donne :

$$\left(\sum_{j=1}^p |\xi_j|^r \right)^{\frac{1-r}{r}} = k^{1-r}.$$

En reportant cette expression dans la première relation, nous obtenons :

$$\alpha_j = \sigma k^{1-r} |\xi_j|^{r-1} \text{signe}(\xi_j)$$

ou encore :

$$\alpha_j \text{signe}(\alpha_j) = \sigma k^{1-r} |\xi_j|^{r-1}.$$

Nous en tirons :

$$\xi_j = \text{signe}(\alpha_j) \sigma^{\frac{1}{1-r}} k |\alpha_j|^{\frac{1}{r-1}}.$$

Nous pouvons donc calculer σ par

$$k = \left(\sum_{j=1}^p |\xi_j|^r \right)^{\frac{1}{r}} = k \sigma^{\frac{1}{1-r}} \left(\sum_{j=1}^p |\alpha_j|^{\frac{r}{r-1}} \right)^{\frac{1}{r}}.$$

Comme $s = \frac{r}{r-1}$, nous obtenons $\sigma = \|\alpha\|_s$.

La preuve se conclut en reprenant l'expression de ξ_j :

$$\xi_j = \text{signe}(\alpha_j) \|\alpha\|_s^{\frac{1}{1-r}} k |\alpha_j|^{\frac{1}{r-1}}$$

et en calculant :

$$\alpha^T \xi = k \left(\sum_{j=1}^p |\alpha_j|^{\frac{1}{r-1}+1} \right) \|\alpha\|_s^{\frac{1}{1-r}} = k \|\alpha\|_s^s \|\alpha\|_s^{1-s} = k \|\alpha\|_s.$$

Les cas $r = 1$ et $r = +\infty$ se prouvent directement ou en passant à la limite. \square

L'application du lemme 3.1 donne

$$\max_x \{ (P^T x)^T \xi \mid \|\xi\|_r \leq k \} = k \|P^T x\|_s.$$

En substituant le maximum calculé au terme $(P\xi)^T x$ du membre de gauche de la contrainte

$$(\bar{a} + P\xi)^T x \leq 0, \quad \forall \xi \in \Xi,$$

nous obtenons la contrainte unique

$$\bar{a}^T x + k \|P^T x\|_s \leq 0$$

qui est l'*équivalent robuste* recherché. Pour les cas $r = 1, 2$ et ∞ considérés dans l'hypothèse 3.2, la contrepartie peut être traitée par des techniques standard de programmation mathématique. Dans ce qui suit, nous notons e tout vecteur dont les composantes sont uniformément égales à 1, de dimension appropriée au contexte.

Cas $r = 1$ La norme conjuguée est ℓ_∞ . La contrepartie robuste

$\bar{a}^T x + k \|P^T x\|_\infty \leq 0$ est équivalente à l'ensemble d'inégalités linéaires

$$\begin{aligned} a^T x + kz &\leq 0 \\ ze &\geq P^T x \\ ze &\geq -P^T x, \end{aligned}$$

où z est une *variable scalaire* et e est le vecteur de composantes 1 de même dimension que le sous-jacent ξ .

Cas $r = \infty$ La norme conjuguée est ℓ_1 . La contrepartie robuste

$\bar{a}^T x + k \|P^T x\|_1 \leq 0$ est équivalente à l'ensemble d'inégalités linéaires

$$\begin{aligned} a^T x + ke^T z &\leq 0 \\ z &\geq P^T x \\ z &\geq -P^T x, \end{aligned}$$

où z est une variable auxiliaire *de même dimension* que le sous-jacent ξ .

Cas $r = 2$ La norme conjuguée est ℓ_2 . La contrepartie robuste

$\bar{a}^T x + k \|P^T x\|_2 \leq 0$ est une contrainte de type conique quadratique. De telles contraintes sont traitées de manière efficace par les méthodes de points intérieurs.

Le calcul de la direction de Cauchy repose sur la dualité en programmation convexe. Nous nous proposons d'établir un résultat semblable dans le cas où l'ensemble d'incertitude est décrit par le polyèdre $\{\xi \mid Q\xi \leq t\}$. Dans ce cas, le test de robustesse pour la contrainte $(\bar{a} + P\xi)^T x \leq 0, \forall \xi \in \Xi$ est donné par $\max_{\xi} \{(P^T x)^T \xi \mid Q\xi \leq t\} \leq -\bar{a}^T x$.

Théorème 3.1 *La condition*

$$\max_{\xi} \{(P^T x)^T \xi \mid Q\xi \leq t\} \leq -\bar{a}^T x$$

est satisfaite si et seulement s'il existe une solution au système linéaire en x et u

$$\begin{aligned} \bar{a}^T x + t^T u &\leq 0 \\ P^T x - Q^T u &= 0 \\ u &\geq 0. \end{aligned}$$

Preuve: Dans la preuve, x est considéré comme un paramètre fixe. Le problème linéaire

$$\max_{\xi} \{(P^T x)^T \xi \mid Q\xi \leq t\}$$

a comme dual

$$\min_u \{t^T u \mid Q^T u - P^T x = 0, u \geq 0\}.$$

Pour toute paire (ξ, u) primale-duale réalisable

$$\left. \begin{aligned} Q\xi &\leq t \\ Q^T u - P^T x &= 0, \quad u \geq 0 \end{aligned} \right\} \Rightarrow (P^T x)^T \xi \leq t^T u,$$

avec égalité si et seulement si la paire (ξ, u) est optimale. Dans le membre de gauche de l'inégalité

$$\bar{a}^T x + \max_{\xi} \{(P^T x)^T \xi \mid Q\xi \leq t\} \leq 0,$$

nous pouvons donc remplacer la quantité maximisée par l'objectif dual $t^T u$ en y ajoutant la contrainte que u doit être dual réalisable. De cette façon, nous obtenons l'équivalent robuste

$$\begin{aligned} \bar{a}^T x + t^T u &\leq 0 \\ P^T x - Q^T u &= 0 \\ u &\geq 0. \end{aligned}$$

□

3.5. PROBABILITÉ DE SATISFACTION DE LA CONTRAINTE POUR UNE SOLUTION ROBUSTE

Il est possible de déduire par le même raisonnement la contrepartie robuste pour les cas d'ensemble d'incertitude sous forme de boule avec $r = 1$ et $r = \infty$. En effet, ces boules sont des polyèdres que nous pouvons décrire explicitement par des inégalités linéaires. Ainsi, $\|\xi\|_\infty \leq k$ est équivalent à

$$-ke \leq \xi \leq ke$$

et $\|\xi\|_1 \leq k$ est équivalent à

$$\begin{aligned}\xi^+ - \xi^- &= \xi \\ e^T(\xi^+ + \xi^-) &\leq k \\ \xi^+ &\geq 0, \xi^- \geq 0.\end{aligned}$$

Cette dernière formulation utilise la décomposition de $\xi = \xi^+ - \xi^-$ en une partie positive et une partie négative.

Interprétation La contrepartie robuste de la contrainte

$$(\bar{a} + P\xi)^T x \leq 0, \forall \xi \text{ tel que } Q\xi \leq t,$$

peut s'interpréter comme une projection. En effet, l'ensemble des solutions robustes x est donc défini par :

$$\{x \mid \exists u \geq 0 \text{ tel que } a^T x + t^T u \leq 0, P^T x - Q^T u = 0\}.$$

Cet ensemble est bien la projection sur l'espace des variables x d'un ensemble d'inégalités en $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^p$.

3.5 Probabilité de satisfaction de la contrainte pour une solution robuste

Nous rappelons rapidement le lien entre l'optimisation robuste et la satisfaction de contrainte en probabilité. Supposons que nous puissions calculer la probabilité que la variable aléatoire ξ appartienne à l'ensemble d'incertitude et que cette probabilité vaille $1 - \alpha$. Nous pouvons alors affirmer qu'imposer que la contrainte soit satisfaite pour tout ξ dans cet ensemble d'incertitude implique que la contrainte sera satisfaite en probabilité avec une probabilité au moins égale à $1 - \alpha$. Nous pouvons nous persuader aisément que la réciproque n'est pas vraie. La formulation par ensemble d'incertitude est donc plus restrictive que la formulation de satisfaction en probabilité.

Nous considérons la contrainte robuste de la forme

$$\begin{aligned}\sum_{i=1}^{n+1} a_i(\xi) x_i &\leq 0 \\ a(\xi) &= \bar{a} + P\xi \\ \xi &\in \Xi \subset \mathbb{R}^p.\end{aligned}$$

L'ensemble Ξ sera pris comme intersection de la boule $B(0, 1)_\infty$ et d'une boule $B(0, k)_r$ avec $r = 1$ ou $r = 2$.

3.5. PROBABILITÉ DE SATISFACTION DE LA CONTRAINTE POUR UNE SOLUTION ROBUSTE

3.5.1 Ensemble d'incertitude $\Xi = B(0, k)_2$

La preuve de la proposition 3.1, à part le lemme 3.3, est empruntée à [3].

Proposition 3.1 Soient ξ_i , $i = 1, \dots, p$ des variables aléatoires indépendantes à valeurs dans l'intervalle $[-1, 1]$ et de moyenne nulle : $E(\xi_i) = 0$. Si z_i , $i = 1, \dots, p$ sont des coefficients déterministes, nous avons pour tout $k \geq 0$

$$\text{Prob} \left\{ \xi \mid \sum_{i=1}^p z_i \xi_i > k \sqrt{\sum_{i=1}^p z_i^2} \right\} \leq \exp\left(-\frac{k^2}{1,5}\right).$$

Remarque 3.1 Le résultat classique donne la borne $\exp(-\frac{k^2}{2})$ qui est plus grande (donc moins bonne).

Les deux lemmes qui suivent seront utilisés dans la démonstration :

Lemme 3.2 (Inégalité de Tchebitcheff)

$$\text{Prob}(X \geq a) \leq e^{-a} E(e^X)$$

Preuve:

$$\begin{aligned} E(e^X) &= E(e^X | X < a) \text{Prob}(X < a) + E(e^X | X \geq a) \text{Prob}(X \geq a) \\ &\geq E(e^X | X \geq a) \text{Prob}(X \geq a) \quad (\text{car } e^X > 0 \Rightarrow E(e^X | X < a) \geq 0) \\ &\geq e^a \text{Prob}(X \geq a). \end{aligned}$$

La dernière inégalité découle de

$$E(e^X | X \geq a) \geq E(e^a | X \geq a) = e^a.$$

□

Lemme 3.3 L'inégalité $e^{\frac{t^2}{\alpha}} + t \geq e^t$ avec $1 \leq \alpha \leq \sqrt{\frac{8}{3}}$ est valide pour tout $t \geq 0$.

Preuve: La fonction $e^{\frac{t^2}{\alpha}} + t - e^t$ est décroissante en α pour tout $t > 0$. Pour prouver la propriété, il suffira de le faire pour la valeur extrême de α . Nous supposons donc dans la suite des calculs que $\alpha = \sqrt{\frac{8}{3}}$. La fonction $g(t) = e^{\frac{t^2}{\alpha}} + t - e^t$ a pour dérivées

$$g'(t) = \frac{2}{\alpha} t e^{\frac{t^2}{\alpha}} + 1 - e^t$$

et

$$g''(t) = \frac{2}{\alpha} \left(\frac{2}{\alpha} t^2 + 1 \right) e^{\frac{t^2}{\alpha}} - e^t = e^t \left(\frac{2}{\alpha} \left(\frac{2}{\alpha} t^2 + 1 \right) e^{\frac{t^2}{\alpha} - t} - 1 \right). \quad (3.9)$$

Nous remarquons que $g'(0) = 0$ et $g''(0) = 2/\alpha - 1 \geq \sqrt{3/2} - 1 > 0$. Montrons d'abord $g''(t) \geq 0$ pour tout $t \geq 0$. Pour cela, il suffit de montrer que le terme entre parenthèses dans le membre de droite de (3.9) est positif. A l'aide de l'inégalité $e^x \geq 1 + x$, nous avons

$$\frac{2}{\alpha} \left(\frac{2}{\alpha} t^2 + 1 \right) e^{\frac{t^2}{\alpha} - t} - 1 \geq h(t) = \frac{2}{\alpha} \left(\frac{2}{\alpha} t^2 + 1 \right) \left(1 + \frac{t^2}{\alpha} - t \right) - 1.$$

3.5. PROBABILITÉ DE SATISFACTION DE LA CONTRAINTE POUR UNE SOLUTION ROBUSTE

Le membre de droite est un polynôme du quatrième degré. Nous vérifions qu'il est convexe car $\alpha < 4$ et

$$h''(t) = \frac{12}{\alpha^2} \left(\frac{4}{\alpha} t^2 - 2t + 1 \right) = \frac{12}{\alpha^2} \left((1-t)^2 + \left(\frac{4}{\alpha} - 1 \right) t^2 \right) > 0.$$

Le polynôme

$$h'(t) = \frac{2}{\alpha} \left(\frac{8}{\alpha^2} t^3 - \frac{6}{\alpha} t^2 + \frac{6}{\alpha} t - 1 \right) = \sqrt{\frac{3}{2}} \left(3t^3 - 3\sqrt{\frac{3}{2}} t^2 + 3\sqrt{\frac{3}{2}} t - 1 \right)$$

est donc croissant ; il prend la valeur $-\sqrt{3/2}$ en 0 et tend vers $+\infty$ lorsque t tend vers $+\infty$. Il a donc une seule racine réelle qui se trouve en $\bar{t} = 0,366$. La fonction $h(t)$ atteint son minimum en \bar{t} et donc $g''(t) \geq h(t) \geq h(\bar{t}) = 0,0208$.

Nous avons donc prouvé que $g''(t) \geq 0$ sur l'intervalle $[0,1]$, et donc que la fonction $g'(t)$ est croissante. Comme $g'(0) = 0$, nous en déduisons que $g'(t) \geq 0$, et donc que $g(t)$ est croissante. Enfin, comme $g(0) = 0$, nous avons $g(t) \geq 0$ sur $0 \leq t \leq 1$, ce qui complète la preuve du lemme. \square

Il est maintenant possible de prouver la proposition 3.1. Le lemme 3.2 permet d'écrire

$$\text{Prob}\left(\sum_i z_i \zeta_i > k \|z\|\right) \leq e^{-k\|z\|} E(e^{\sum_i z_i \zeta_i}).$$

Comme les variables ζ_i sont indépendantes, alors $E(e^{\sum_i z_i \zeta_i}) = E(\prod_i e^{z_i \zeta_i}) = \prod_i E(e^{z_i \zeta_i})$. Or

$$\begin{aligned} E(e^{z_i \zeta_i}) &= 1 + \sum_{k \geq 2} E\left(\frac{(z_i \zeta_i)^k}{k!}\right) \quad (\text{car } E(\zeta_i) = 0) \\ &\leq 1 + \sum_{k \geq 2} \frac{|z_i|^k}{k!} \quad (\text{car } |z_i \zeta_i| \leq |z_i|) \\ &\leq e^{|z_i|} - |z_i| \leq e^{\frac{z_i^2}{\alpha}}. \end{aligned}$$

La dernière inégalité découle du lemme 3.3. Nous pouvons donc écrire

$$\text{Prob}\left(\sum_i z_i \zeta_i > k \|z\|\right) \leq e^{-k\|z\|} \prod_i e^{\frac{z_i^2}{\alpha}} = e^{-k\|z\| + \frac{\|z\|^2}{\alpha}}.$$

Comme la relation

$$\text{Prob}\left(\sum_i z_i \zeta_i > k \|z\|\right) = \text{Prob}\left(\gamma \sum_i z_i \zeta_i > \gamma k \|z\|\right)$$

est vraie pour tout $\gamma > 0$, nous avons donc

$$\begin{aligned} \text{Prob}\left(\sum_i z_i \zeta_i > k \|z\|\right) &\leq \min_{\gamma > 0} e^{-\gamma k \|z\| + \gamma^2 \frac{\|z\|^2}{\alpha}} \\ &\leq e^{-\frac{\alpha^2 k^2}{4}} \leq e^{-\frac{k^2}{1,5}}. \end{aligned}$$

3.5. PROBABILITÉ DE SATISFACTION DE LA CONTRAINTE POUR UNE SOLUTION ROBUSTE

□

Une étude expérimentale de la fonction $e^{\frac{t^2}{\alpha}} + t - e^t$ montre que la valeur maximum de α pour laquelle le lemme 3.3 reste vrai est plus grande et vaut 1,79. La borne dans la proposition devient alors $e^{-\frac{t^2}{1,25}}$.

Appliquons le résultat lorsque nous choisissons comme ensemble d'incertitude la boule $B(0, k)_2$ et que nous savons par ailleurs que le support des variables ξ_j est l'intervalle $[-1, 1]$. Considérons la contrainte :

$$\bar{a}^T x + (P\xi)^T x \leq 0, \forall \xi \in B(0, k)_2.$$

D'après le lemme 3.1,

$$\max_{\xi \in \Xi} (P^T x)^T \xi \leq \max_{\xi} \{(P^T x)^T \xi \mid \|\xi\|_2 \leq k\} = k \|P^T x\|_2.$$

En posant $z = P^T x$, la proposition 3.1 donne

$$\text{Prob} \{(\bar{a} + P\xi)^T x \leq 0 \mid \xi \in \Xi\} \geq \text{Prob} \{(P^T x)^T \xi \leq k \|P^T x\|_2\} \geq 1 - \exp\left(-\frac{k^2}{1,5}\right).$$

Nous notons qu'avec $k = 2,63$ nous nous garantissons une probabilité égale à 0,99.

Remarque 3.2 *Le résultat est établi pour $\Xi = B(0, k)_2$ lorsque nous savons par ailleurs que l'hypothèse $\|\xi\|_\infty \leq 1$ est satisfaite. Si, sous cette même hypothèse, nous choisissons l'ensemble d'incertitude plus petit $\Xi = B(0, 1)_\infty \cap B(0, k)_2$, la probabilité $\text{Prob} \{(\bar{a} + \tilde{P}\xi)^T x \leq 0 \mid \tilde{\xi} \in \Xi\}$ est plus faible ; l'inégalité du théorème est donc vraie a fortiori.*

3.5.2 Ensemble d'incertitude $\Xi = B(0, 1)_\infty \cap B(0, k)_1$

Par un corollaire du théorème précédent nous pouvons borner la probabilité associée à un ensemble d'incertitude défini dans les normes ℓ_1 et ℓ_∞ .

Corollaire 3.1 (de la proposition 3.1) *Soient $\xi_i, i = 1, \dots, p$, des variables aléatoires indépendantes à valeurs dans l'intervalle $[-1, 1]$ et de moyenne nulle : $E(\xi_i) = 0$. Si $z_i, i = 1, \dots, p$, sont des coefficients déterministes, nous avons pour tout $k \geq 0$:*

$$\text{Prob} \left\{ \xi \mid \sum_{i=1}^p z_i \xi_i > k \sum_{i=1}^p |z_i| \right\} \leq \exp\left(-\frac{k^2}{1,5p}\right).$$

Preuve: La preuve découle de la relation entre les normes ℓ_1 et ℓ_2 dans \mathbb{R}^p qui implique l'inclusion

$$B(0, k/\sqrt{p})_2 \subseteq B(0, k)_1.$$

Il suffit de remplacer k par k/\sqrt{p} dans la proposition 3.1 pour obtenir le résultat. □

3.6 Eléments pour définir l'ensemble d'incertitude

Le choix d'un ensemble d'incertitude est crucial dans la mise en œuvre de la programmation robuste. Sa forme et sa dimension sont des données arbitraires, laissées à l'appréciation du décideur, qu'il faut spécifier dans la mise en œuvre. Nous discutons ici le cas d'une contrainte robuste de la forme

$$(\bar{a} + P\xi)^T x \leq 0$$

où ξ est un aléa qui appartient à une boule ℓ_r ($r = 1, 2$ ou ∞) ou à l'intersection de deux boules ℓ_1 et ℓ_∞ de rayon déterminé. Bien que la programmation robuste soit souvent mise en avant comme une façon de gérer des problèmes d'optimisation dans l'incertain sans référence à des probabilités, nous nous appuyons sur la similitude entre contre-partie robuste et satisfaction de contrainte en probabilité pour tirer un argument heuristique de choix des paramètres, en particulier du rayon des boules à prendre en considération.

3.6.1 Une approche heuristique basée sur un aléa gaussien

Partons de l'hypothèse que le vecteur de paramètres incertains $a(\xi)$ suit une loi normale multivariée. Dans ce cas, nous pouvons écrire $a(\xi) = \bar{a} + P\xi$, où $\xi = (\xi_1, \dots, \xi_p)$ est un vecteur de variables aléatoires indépendantes identiquement distribuées selon la loi normale standard (centrée réduite). Sous cette hypothèse, il est possible d'assigner une probabilité à l'ensemble d'incertitude défini par la boule $\ell_2 : \|\xi\|_2 \leq k$. En effet, la variable $t = \sum_{i=1}^p \xi_i^2$ suit une loi de χ^2 à p degrés de liberté. Cette loi est tabulée et permet de calculer

$$\text{Prob}(\tilde{\xi} \mid \|\xi\|_2 \leq k) = \text{Prob}(t \leq k^2)$$

et inversement de déterminer k tel que

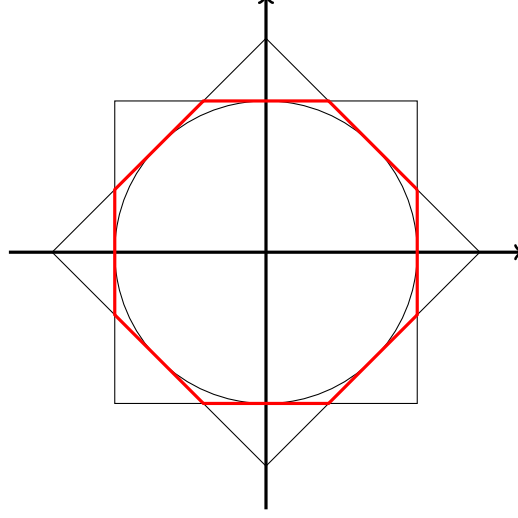
$$\text{Prob}(\tilde{\xi} \mid \|\xi\|_2 \leq k) = 1 - \alpha.$$

Si nous désirons rester dans le domaine de la programmation linéaire, nous pouvons essayer d'approximer la boule ℓ_2 par un polyèdre régulier. Malheureusement, le nombre de facettes nécessaire pour réaliser un niveau d'approximation donné croît exponentiellement avec ce niveau. Dans [10], Ben-Tal et Nemirovski ont proposé une ingénieuse technique d'immersion/projection qui garantit une approximation polynomiale. Nous renvoyons le lecteur à cet article et nous nous contenterons ici de considérer le cas beaucoup plus simple d'une approximation de la boule dans ℓ_2 par l'intersection d'une boule ℓ_1 et d'une boule ℓ_∞ . Cette approche est beaucoup plus fruste, mais elle est commode et donne des résultats tout à fait acceptables en pratique.

Nous pouvons maintenant définir l'ensemble d'incertitude :

$$\Xi = \{\xi : \sum_{i=1}^p |\xi_i| \leq k_1, \max_{i=1, \dots, p} |\xi_i| \leq k_2\}. \quad (3.10)$$

Le choix des paramètres k , k_1 et k_2 est libre. Si nous pensons qu'une boule dans ℓ_2 de rayon k donne une représentation appropriée de l'ensemble d'incertitude, nous pouvons penser prendre


 FIG. 3.3 – Représentation 2D de l'ensemble $\ell_1 \cap \ell_\infty$

$k_1 = k\sqrt{p}$ et $k_2 = k$. Nous obtenons ainsi une approximation extérieure de la boule dans ℓ_2 , comme le montre la figure 3.3. Cette approximation conduit à une solution plus conservative, puisque l'intersection des deux boules ℓ_1 et ℓ_∞ contient la boule ℓ_2 . Lorsque la dimension p croît, cet écart d'approximation peut se révéler excessif. Dans ce cas, nous prendrons $k_1 = \sigma\sqrt{p}$ et $k_2 = \sigma$, avec $\sigma < k$ tel que le volume de l'intersection soit égal à celui de la boule ℓ_2 de rayon k .

3.6.2 Equivalent robuste

Supposons que l'ensemble d'incertitude Ξ soit l'intersection de deux boules ℓ_1 et ℓ_∞ de rayons respectifs k_1 et k_2 . La représentation des deux boules ℓ_1 et ℓ_∞ de l'ensemble d'incertitude (3.10) par des inégalités fait appel à des valeurs absolues. Pour entrer dans le cadre de la programmation linéaire, nous remplaçons les variables ξ_i par une décomposition en une partie positive $\xi_i^+ = \max\{\xi_i, 0\}$ et une partie négative $\xi_i^- = \max\{-\xi_i, 0\}$, soit

$$\xi_i = \xi_i^+ - \xi_i^- \quad \text{et} \quad \xi_i^+ \xi_i^- = 0, \quad i = 1, \dots, p.$$

L'ensemble Ξ est donc décrit par les inégalités

$$\begin{aligned} \xi_i &= \xi_i^+ - \xi_i^-, \quad i = 1, \dots, p \\ \sum_{i=1}^p (\xi_i^+ + \xi_i^-) &\leq k_1 \\ 0 &\leq \xi_i^+ \leq k_2, \quad i = 1, \dots, p \\ 0 &\leq \xi_i^- \leq k_2, \quad i = 1, \dots, p \\ \xi_i^+ \xi_i^- &= 0, \quad i = 1, \dots, p. \end{aligned} \tag{3.11}$$

Pour tester si la contrainte

$$\bar{a}^T x + (P\xi)^T x \leq 0, \quad \forall \xi \in \Xi,$$

pour un x donné, est satisfaite, nous considérons le problème d'optimisation

$$\begin{aligned} \max_{\xi^+, \xi^-} \quad & \sum_{i=1}^p (\xi_i^+ (P^T x)_i - \xi_i^- (P^T x)_i) \\ & \sum_{i=1}^n (\xi_i^+ + \xi_i^-) \leq k_1 \\ & 0 \leq \xi_i^+ \leq k_2, i = 1, \dots, p \\ & 0 \leq \xi_i^- \leq k_2, i = 1, \dots, p \\ & \xi_i^+ \xi_i^- = 0, i = 1, \dots, p. \end{aligned}$$

Il est possible d'éliminer la contrainte de complémentarité $\xi_i^+ \xi_i^- = 0$ sans modifier la valeur de la fonction objectif. En effet, si $\bar{\xi}$ est une solution réalisable avec $\bar{\xi}_i^+ \bar{\xi}_i^- > 0$, pour un i donné, la solution $\xi_i^+ = \bar{\xi}_i^+ - \min\{\bar{\xi}_i^+, \bar{\xi}_i^-\}$ et $\xi_i^- = \bar{\xi}_i^- - \min\{\bar{\xi}_i^+, \bar{\xi}_i^-\}$, pour tout i satisfait toutes les contraintes, y compris la contrainte de complémentarité, alors que l'objectif reste inchangé, puisque $\xi = \bar{\xi}$.

Sans la contrainte de complémentarité, le problème est linéaire. Nous pouvons écrire son dual

$$\begin{aligned} \min_{v, u^+, u^-} \quad & k_1 v + k_2 \left(\sum_{i=1}^p u_i^+ + \sum_{i=1}^p u_i^- \right) \\ & v + u_i^+ \geq (P^T x)_i, i = 1, \dots, p \\ & v + u_i^- \geq -(P^T x)_i, i = 1, \dots, p \\ & v \geq 0, u_i^+ \geq 0, u_i^- \geq 0, i = 1, \dots, p. \end{aligned}$$

D'après le théorème 3.1, l'équivalent robuste s'écrit

$$\begin{aligned} \bar{a}^T x + k_1 v + k_2 (\sum_{i=1}^p u_i^+ + \sum_{i=1}^p u_i^-) & \leq 0 \\ v + u_i^+ & \geq (P^T x)_i, i = 1, \dots, p \\ v + u_i^- & \geq -(P^T x)_i, i = 1, \dots, p \\ v \geq 0, u_i^+ & \geq 0, u_i^- \geq 0, i = 1, \dots, p. \end{aligned} \tag{3.12}$$

L'équivalent robuste (3.12) est linéaire en x , u_i^+ , u_i^- et v . Le nouveau système d'inégalités comprend $2p$ contraintes supplémentaires et $2p+1$ variables non-négatives. Cet accroissement est en partie dû au choix de la norme $\ell_1 \cap \ell_\infty$, et, en particulier, l'utilisation de la norme ℓ_∞ . En effet, si nous ne considérons qu'une norme ℓ_∞ lors de la construction de l'ensemble d'incertitude, l'équivalent robuste rajoute $2p$ variables auxiliaires non négatives et un facteur $2p$ pour les contraintes, soit quasiment le même nombre de variables et de contraintes qu'avec l'ensemble $\ell_1 \cap \ell_\infty$. Le raffinement avec ℓ_1 ne rajoute au final qu'une variable auxiliaire non négative supplémentaire.

Remarque 3.3 *Il est parfois possible de déterminer a priori le signe des composants du vecteur $P^T x$. Dans ce cas, nous pouvons supprimer dans le système d'inégalités linéaires (3.12) p contraintes et p variables. En effet, si $(P^T x)_i \geq 0$, alors $v + u_i^- \geq -(P^T x)_i$ est redondant puisque $v \geq 0$ et $u_i^- \geq 0$. De même, si $(P^T x)_i < 0$ nous éliminerons la contrainte $v + u_i^+ \geq (P^T x)_i$.*

Citons deux conditions suffisantes d'application de la propriété énoncée. Dans les deux cas, nous supposons $x \geq 0$. Un premier exemple de validité de la propriété est donnée par $P \geq 0$ (ou $P \leq 0$). Le deuxième exemple est celui d'une matrice diagonale P . Nous avons alors que le signe de $(P^T x)_i$ est celui de l'élément diagonal P_{ii} .

3.7 Problème général d'optimisation linéaire

Considérons le problème de programmation linéaire avec données incertaines qui a été introduit dans les premières sections :

$$\begin{aligned} \min_x \quad & \sum_{j=1}^n c_j(\xi) x_j \\ & \sum_{j=1}^n a_{ij}(\xi) x_j \leq b_i(\xi), \quad i = 1, \dots, m_1 \\ & \sum_{j=1}^n g_{ij} x_j = h_i, \quad i = m_1 + 1, \dots, m_2. \end{aligned}$$

La contrepartie robuste de ce problème pour un ensemble d'incertitude Ξ a été formulée à l'aide d'une variable auxiliaire z comme

$$\min_{x,z} \quad z \tag{3.13a}$$

$$\sum_{j=1}^n c_j(\xi) x_j - z \leq 0, \quad \forall \xi \in \Xi \tag{3.13b}$$

$$\sum_{j=1}^n a_{ij}(\xi) x_j \leq b_i(\xi), \quad \forall \xi \in \Xi, \quad i = 1, \dots, m_1 \tag{3.13c}$$

$$\sum_{j=1}^n g_{ij} x_j = h_i, \quad \forall \xi \in \Xi, \quad i = m_1 + 1, \dots, m_2. \tag{3.13d}$$

Jusqu'à présent, nous avons exclu les contraintes d'égalité. Il faut donc discuter s'il est possible de les prendre en charge et à quelles conditions. L'autre question à traiter dans cette section concerne l'objectif. Devons-nous nécessairement nous intéresser à minimiser la pire des performances possible? Ne devrions-nous pas prendre en compte la performance moyenne, et, dans l'affirmative, comment le faire?

3.7.1 Equivalent robuste d'une contrainte d'égalité

En général nous ne considérons pas le cas des contraintes d'égalité en programmation robuste. En effet une contrainte d'égalité implique que toute variation de l'aléa n'altère pas l'égalité. Si l'ensemble d'incertitude sur l'aléa a un intérieur non vide, la condition sur l'égalité se transforme

en un ensemble d'identités qui, sauf dégénérescence, sont si restrictives qu'elles ne laissent pas la place à une solution réalisable.

Prenons l'exemple d'une contrainte $g^T x = h$, dans laquelle les coefficients g et h sont incertains. Supposons que nous souhaitons rendre la contrainte robuste vis-à-vis de tous les aléas décrits par l'ensemble d'incertitude

$$\begin{aligned} g(\xi) &= \bar{g} + G\xi \\ h(\xi) &= \bar{h} + h^T \xi \\ \xi &\in \Xi. \end{aligned}$$

(Dans cet exemple, $g \in \mathbb{R}^n, G \in \mathbb{R}^{n \times p}, \bar{g} \in \mathbb{R}$ et $h \in \mathbb{R}^p$.) Nous faisons de plus l'hypothèse raisonnable que Ξ a un intérieur non vide qui contient l'origine (nous pouvons nous ramener à ce cas quand Ξ est convexe non réduit à un point). Il est facile de voir que pour que l'égalité

$$(\bar{g} + G\xi)^T x = \bar{h} + h^T \xi$$

soit vraie pour tout $\xi \in \Xi$, il faut que

$$\bar{g}^T x = \bar{h} \tag{3.14a}$$

$$G^T x = h. \tag{3.14b}$$

La contrepartie robuste d'une contrainte d'égalité est le système (3.14) de $p + 1$ équations à n inconnues. Il est sur-déterminé si $p + 1 > n$ et peut ne pas avoir de solution. Et même si $p + 1 \leq n$, il peut restreindre excessivement l'ensemble des solutions réalisables x et ne plus être compatible avec les autres restrictions du problème. Il existe cependant des cas d'application où cette approche est justifiée et conduit à des résultats intéressants [34].

3.7.2 Performance moyenne et performance garantie

La valeur optimale z^* du problème (3.13) s'interprète comme un majorant de la fonction économique pour toute réalisation de l'aléa dans l'ensemble d'incertitude. Cette solution ne dit en revanche rien sur la performance nominale, celle que nous pourrions observer en moyenne dans une simulation où ξ serait tiré aléatoirement suivant une distribution compatible avec les hypothèses faites sur l'ensemble d'incertitude Ξ . Dans certaines applications, particulièrement celles liées aux problèmes multipériodiques, nous constatons souvent que le coût

$$z(\xi) = \sum_{j=1}^n c_j(\xi) x_j^*,$$

pour une solution optimale x_j^* , possède, sur un échantillon aléatoire de ξ , une distribution très resserrée autour d'une valeur moyenne proche du « pire cas » z^* .

Dans la pratique, le critère de performance moyenne est indéniablement pertinent. Il est donc intéressant d'essayer de le prendre en compte, fut-ce au prix d'une certaine détérioration de la valeur du « pire cas ». Mais que peut-être la « valeur moyenne » de l'objectif ? Faisons l'hypothèse

que l'ensemble d'incertitude Ξ est symétrique et que son centre est l'origine et coïncide avec la moyenne de ξ . Dans ce cas, nous pouvons considérer que

$$\hat{z} = \sum_{j=1}^n c_j(0)x_j^*$$

est un estimateur de la valeur moyenne de l'objectif. Le problème paramétrique en γ :

$$\begin{aligned} \min_x \quad & \sum_{j=1}^n c_j(0)x_j^* \\ & \sum_{j=1}^n c_j(\xi)x_j \leq z^* + \gamma, \quad \forall \xi \in \Xi \\ & \sum_{j=1}^n a_{ij}(\xi)x_j \leq b_i(\xi), \quad \forall \xi \in \Xi, \quad i = 1, \dots, m_1. \\ & \sum_{j=1}^n g_{ij}x_j = h_i, \quad \forall \xi \in \Xi, \quad i = m_1 + 1, \dots, m_2, \end{aligned}$$

avec $\gamma > 0$ permet de réaliser un arbitrage entre la performance moyenne et la performance garantie dans le pire cas.

3.8 Problèmes dynamiques

La présence de plusieurs étapes dans un problème d'optimisation apporte une difficulté conceptuelle nouvelle. Dans un modèle déterministe, il n'y a pas vraiment lieu de distinguer un ordre temporel dans l'algorithme de résolution. Bien sûr, l'évolution est déterminée par des équations de dynamique liant des variables d'indices temporels différents, mais il n'y a pas d'obligation de tenir compte de cet ordre dans l'agencement des calculs. Dans un environnement incertain, la situation est totalement différente. En général, l'aléa de première étape est révélé au moment de la prise de décision de la deuxième étape. Ne pas en tenir compte priverait le décideur d'une faculté d'adaptation à la réalisation des aléas. Il est donc normal de définir les variables de décision de deuxième étape comme des fonctions de l'aléa de première étape. Cela veut dire, remplacer des variables réelles de dimension finie par des fonctions, soit un pas vers un ordre de complexité plus élevé. La difficulté ne s'arrête pas là. A la première étape, les décisions devront tenir compte de cet état des choses, c'est-à-dire anticiper les conséquences immédiates de l'aléa de première étape, mais aussi anticiper la fonction représentant la décision en deuxième étape. Une solution à un problème d'optimisation stochastique à plusieurs étapes est donc d'un ordre de grandeur très supérieur à ce que nous avons traité précédemment. Il n'est pas directement traitable par une approche robuste.

Pour avoir une chance de proposer une solution opérationnelle à un tel problème, il faut en réduire la complexité. Une piste possible consiste à réduire l'espace fonctionnel des décisions de deuxième période. Nous parlons alors de règle de décision, concept qui a été introduit dans le

cadre d'un problème de commande [29], et dans celui de la programmation stochastique ([26] ou encore [22]). Lorsque les décisions de deuxième période sont restreintes à des fonctions affines de l'aléa de première période, nous parlons de règle de décision linéaire. Cette restriction restreint le champ du possible et génère des solutions sous-optimales, mais ces solutions sont obtenues en résolvant des problèmes d'optimisation incomparablement plus simples. Le concept de règles de décision a été repris en optimisation robuste sous le vocable AARC (Affinely Adjustable Robust Counterpart) dans l'article [6].

La formalisation de cette idée s'effectue comme suit. Considérons la contrainte type d'un modèle à deux étapes :

$$a_1(\xi_1)^T x_1 + a_2(\xi_1, \xi_2)^T x_2 \leq b(\xi_1, \xi_2).$$

Pour fixer les idées, nous supposons que x_1 et x_2 sont des variables de dimension n et m respectivement. De même les coefficients a_1 et a_2 sont des vecteurs de dimension n et m respectivement. La dépendance des coefficients a_1 , a_2 et b vis-à-vis de l'aléa sous-jacent (ξ_1, ξ_2) s'exprime par les fonctions affines

$$a_1(\xi_1) = \bar{a}_1 + P_1 \xi_1 \quad (3.15a)$$

$$a_2(\xi_1, \xi_2) = \bar{a}_2 + P_{21} \xi_1 + P_{22} \xi_2 \quad (3.15b)$$

$$b(\xi_1, \xi_2) = \bar{b} + b_1^T \xi_1 + b_2^T \xi_2. \quad (3.15c)$$

Dans cette dernière relation, b_1 et b_2 sont des vecteurs de mêmes dimensions que ξ_1 et ξ_2 . L'ensemble d'incertitude Ξ est supposé satisfaire l'hypothèse 3.2 et donc être polyédral.

Le caractère adaptatif de la variable x_2 est pris en compte par une fonction affine de ξ_1

$$x_2(\xi_1) = \bar{x}_2 + D \xi_1.$$

Dans cette expression, les composantes du vecteur \bar{x}_2 et de la matrice D sont les variables de décision de cette nouvelle formulation. Ces variables sont à déterminer antérieurement à la réalisation de l'aléa ξ_1 . Le caractère affine est arbitraire. Cependant, il a été introduit comme tel dans [6] avec des résultats satisfaisants. D'autre part, il permet de conserver la linéarité de l'équivalent robuste.

L'équivalent robuste de la contrainte dans ce problème à deux étapes est donc

$$(\bar{a}_1 + P_1 \xi_1)^T x_1 + (\bar{a}_2 + P_{21} \xi_1 + P_{22} \xi_2)^T (\bar{x}_2 + D \xi_1) - (\bar{b} + b_1 \xi_1 + b_2 \xi_2) \leq 0, \forall (\xi_1, \xi_2) \in \Xi.$$

Pour construire la contrepartie robuste, il faut résoudre le problème d'optimisation

$$\begin{aligned} \max_{\xi_1, \xi_2} \quad & (P_1^T x_1 + P_{21}^T \bar{x}_2 + D^T \bar{a}_2 - b_1) \xi_1 + (P_{22}^T \bar{x}_2 - b_2)^T \xi_2 + (P_{21} \xi_1 + P_{22} \xi_2)^T D \xi_1 \\ & (\xi_1, \xi_2) \in \Xi, \end{aligned} \quad (3.16)$$

dans lequel les variables x_1 , x_2 et D sont fixées et traitées comme des paramètres. Nous réalisons que ce problème est quadratique et fait intervenir les termes du second ordre

$$(P_{21} \xi_1 + P_{22} \xi_2)^T D \xi_1.$$

3.8.1 Problèmes avec recours fixe

Pour rester dans le cadre de la programmation linéaire, nous introduisons l'hypothèse suivante.

Hypothèse 3.3 *Le vecteur $a_2(\xi)$ associé au recours est fixe ($a_2(\xi) \equiv a_2$). Nous parlons alors de « recours fixe ».*

Théorème 3.2 *Soit un problème dynamique avec recours fixe et un ensemble d'incertitude polyédral. Si nous adoptons une règle de décision linéaire, l'équivalent robuste d'une contrainte dynamique est un ensemble défini par un nombre fini d'inégalités linéaires.*

Preuve: L'hypothèse de recours fixe signifie que $P_{21}\xi_1 + P_{22}\xi_2 \equiv 0$. Le terme quadratique dans l'objectif du problème d'optimisation permettant de construire l'équivalent robuste disparaît. Nous nous retrouvons donc dans le cadre linéaire usuel. \square

L'implication pratique de l'hypothèse de recours fixe est que l'aléa de deuxième période est confiné au membre de droite de la contrainte. Ceci exclut notamment la gestion des aléas de prix de seconde période. Cette hypothèse est donc restrictive, mais elle permet tout de même de considérer des cas pertinents comme ceux d'un aléa sur la demande dans un problème de production ou de chaîne d'approvisionnement [5], ou un aléa sur les apports d'eau dans un problème de gestion de barrages hydrauliques [2].

3.8.2 Règle de décision linéaire par morceaux avec coefficients de recours fixes

La présente section propose une extension de la règle de décision linéaire qui différencie le recours en fonction du signe de l'aléa. Cette extension s'applique au cas d'un coefficient de recours fixe. Pour simplifier la présentation, nous considérons un problème à deux étapes avec comme contrainte typique

$$(\bar{a}_1 + P_1\xi_1)^T x_1 + a_2^T x_2(\xi_1) \leq \bar{b} + b_1^T \xi_1 + b_2^T \xi_2.$$

La règle proposée s'écrit

$$x_2(\xi_1) = \bar{x}_2 + D\xi_1 + D_p \max\{\xi_1, 0\} + D_n \max\{-\xi_1, 0\}.$$

Dans cette expression, le vecteur $\max\{\xi_1, 0\}$ est obtenu en appliquant l'opérateur de maximisation, coordonnée par coordonnée :

$$(\max\{\xi_1, 0\})_i = \max\{\xi_{1i}, 0\}, \quad i = 1, \dots, p.$$

Les éléments à choisir dans cette règle de décision sont, outre le vecteur \bar{x}_2 et la matrice $D \in \mathbb{R}^{n \times p}$, les deux matrices D_p et D_n , toutes deux dans $\mathbb{R}^{n \times p}$.

La condition de robustesse sur la contrainte avec la règle de décision choisie est donc :

$$(\bar{a}_1 + P_1\xi_1)^T x_1 + a_2^T [\bar{x}_2 + D\xi_1 + D_p \max\{\xi_1, 0\} + D_n \max\{-\xi_1, 0\}] \leq \bar{b} + b_1^T \xi_1 + b_2^T \xi_2, \quad \forall (\xi_1, \xi_2) \in \Xi.$$

3.8. PROBLÈMES DYNAMIQUES

L'ensemble Ξ est un polyèdre quelconque. Pour formuler la condition de robustesse, nous considérons le problème d'optimisation auxiliaire :

$$\max_{(\xi_1, \xi_2) \in \Xi} (P_1^T x_1 + D^T a_2 - b_1)^T \xi_1 + (D_p^T a_2)^T \max\{\xi_1, 0\} + (D_n^T a_2)^T \max\{-\xi_1, 0\} - b_2^T \xi_2. \quad (3.17)$$

Les fonctions $\max\{\xi_1, 0\}$ et $\max\{-\xi_1, 0\}$ sont convexes. Une condition suffisante pour que la fonction objectif soit concave est que les coefficients des opérateurs max soient négatifs, soit

$$D_n^T a_2 \leq 0 \quad \text{et} \quad D_p^T a_2 \leq 0.$$

Nous nous proposons de montrer que, sous la condition énoncée, nous pouvons formuler le problème d'optimisation auxiliaire (3.17) comme le problème de programmation linéaire suivant

$$\begin{aligned} \max_{\xi_1^+, \xi_1^-, \xi_2} \quad & (P_1^T x_1 + D^T a_2 - b_1)^T \xi_1 + (D_p^T a_2)^T \xi_1^+ + (D_n^T a_2)^T \xi_1^- - b_2^T \xi_2 \\ & \xi_1 = \xi_1^+ - \xi_1^- \\ & \xi_1^+ \geq 0, \xi_1^- \geq 0 \\ & (\xi_1, \xi_2) \in \Xi. \end{aligned}$$

La preuve de la proposition s'obtient par un raisonnement semblable à celui utilisé dans la section 3.6.2. Il est facile de voir que pour toute solution réalisable $(\xi_1, \xi_1^+, \xi_1^-)$, le triplet $(\xi_1, \max\{\xi_1, 0\}, \max\{-\xi_1, 0\})$ est réalisable et améliore la fonction objectif. Il suffit de remarquer que $\xi_1^+ - \min\{\xi_1^+, \xi_1^-\} = \max\{\xi_1, 0\}$ et $\xi_1^- - \min\{\xi_1^+, \xi_1^-\} = \max\{-\xi_1, 0\}$. Ce changement de variable laisse l'objectif inchangé à l'exception du terme supplémentaire

$$-(D_p^T a_2 + D_n^T a_2)^T \min\{\xi_1^+, \xi_1^-\} \geq 0.$$

La règle de décision linéaire par morceaux s'écrit donc

$$x_2(\xi_1) = \bar{x}_2 + D\xi_1 + D_p \max\{\xi_1, 0\} + D_n \max\{-\xi_1, 0\},$$

avec comme contrainte sur les matrices D_p et D_n

$$D_n^T a_2^j \leq 0 \quad \text{et} \quad D_p^T a_2^j \leq 0$$

pour toute contrainte j où intervient un recours.

3.8.3 Problèmes avec coefficient de recours incertain

Dans le cas d'un recours avec coefficients incertains, les matrices P_{21} et P_{22} ne sont pas nulles. Comme D est une variable de décision non contrainte, la forme quadratique dans l'objectif du problème (3.16) peut être indéfinie. Lorsque l'ensemble Ξ est un polyèdre, nous ne pouvons pas construire un problème dual ayant même valeur optimale que le problème (3.16). En revanche, si Ξ est un ellipsoïde, le problème (3.16) est équivalent à un problème convexe sur le cône des matrices semi-définies positives. Nous ne développons pas cette généralisation dans le cadre restreint

de la programmation linéaire adopté dans la présente introduction : le lecteur se reportera au livre [9] et à l'article [6] pour plus de détails. Le point essentiel que le lecteur doit réaliser est que les codes disponibles en programmation conique, tel Sedumi [39], permettent de résoudre avec très grande efficacité des problèmes de taille respectable, ce qui rend la généralisation discutée tout à fait opérationnelle.

3.9 Extension du concept de robustesse : violation contrôlée des contraintes

La programmation robuste ne prend en considération que les aléas qui font partie de l'ensemble d'incertitude. Une solution «robuste» peut donc violer les contraintes pour certains aléas qui ne font pas partie de l'ensemble d'aléas, à un niveau éventuellement inacceptable. Nous pourrions songer à accroître l'ensemble d'incertitude pour y inclure une plus grande proportion d'événements rares, mais ce n'est pas une réponse adéquate au problème. En effet, en agrandissant l'ensemble d'incertitude, nous diminuons l'ensemble des solutions robustes, au point que cet ensemble peut devenir vide. De fait il est généralement impossible qu'une même solution respecte la contrainte dans la totalité des cas de figure. Il faut donc admettre qu'une violation est possible. Mais dans quelle mesure ?

Dans [4], les auteurs proposent une réponse qui s'appuie sur l'idée suivante. Une solution *robuste étendue* doit satisfaire deux critères :

1. Elle doit être robuste au sens habituel, c'est-à-dire que pour toute réalisation de l'aléa dans l'ensemble d'incertitude, la contrainte doit être exactement satisfaite.
2. Si l'aléa n'appartient pas à l'ensemble d'incertitude, une violation de la contrainte peut être acceptée, mais cette violation doit être «contrôlée».

Par définition, une solution robuste étendue est donc robuste au sens classique. Le concept de solution robuste étendue est donc plus restrictif, mais il couvre tous les cas de figure de l'incertitude.

Que faut-il entendre par «violation contrôlée» ? Les auteurs de [4] suggèrent que la violation soit bornée par une fonction proportionnelle à la distance de l'aléa à l'ensemble d'incertitude. Sans avoir encore défini cette distance, nous voyons tout de suite que, si l'aléa appartient à l'ensemble d'incertitude, la violation doit être nulle puisque la distance à l'ensemble d'incertitude est nulle. Nous couvrons ainsi les deux conditions dans la définition d'une solution étendue.

3.9.1 Définition formelle de la robustesse étendue

Pour formaliser l'idée, il faut définir une «distance» à l'ensemble d'incertitude. Faisons-le d'une manière abstraite pour commencer. Comme dans nos exemples précédents, la contrainte

$$(\bar{a} + P\xi)^T x = \bar{a}^T x + (P^T x)^T \xi \leq 0$$

doit être satisfaite pour tous les ξ dans l'ensemble d'incertitude.

3.9. EXTENSION DU CONCEPT DE ROBUSTESSE : VIOLATION CONTRÔLÉE DES CONTRAINTES

Nous supposons que $\Omega = \mathbb{R}^p$ et que Ξ est convexe. Nous introduisons une fonction distance convexe $d(\xi, \Xi)$ entre ξ et l'ensemble Ξ . Une solution sera dite *robuste étendue* si elle satisfait

$$\bar{a}^T x + (P^T x)^T \xi \leq \alpha d(\xi, \Xi), \forall \xi \in \Omega$$

où $\alpha > 0$ est un paramètre choisi par le décideur qui contrôle l'ampleur de la violation tolérée. Notez que ξ est pris dans tout l'ensemble des possibles. Nous testons la robustesse étendue d'une solution x en vérifiant la condition

$$\max_{\xi \in \Omega} \{(P^T x)^T \xi - \alpha d(\xi, \Xi)\} \leq -\bar{a}^T x.$$

Soit maintenant $\delta : \mathbb{R}^p \rightarrow \mathbb{R}_+$ une norme sous-tendant la distance d . Ainsi

$$d(\xi, \Xi) = \min_{\xi' \in \Xi} \delta(\xi - \xi').$$

La contrepartie robuste étendue de la contrainte s'obtient en résolvant le problème d'optimisation convexe

$$\begin{aligned} & \max_{\xi \in \mathbb{R}^p} \{(P^T x)^T \xi - \alpha \min_{\xi' \in \Xi} \delta(\xi - \xi')\} \\ &= \max_{\xi \in \mathbb{R}^p} \{(P^T x)^T \xi + \alpha \max_{\xi' \in \Xi} \{-\delta(\xi - \xi')\}\} \\ &= \max \{(P^T x)^T \xi - \alpha \delta(\xi - \xi') \mid \xi \in \mathbb{R}^p, \xi' \in \Xi\}. \end{aligned}$$

3.9.2 Robustesse étendue dans le cadre de la programmation linéaire

Comme dans les sections précédentes, nous voulons construire une contre-partie robuste de l'inégalité de départ en termes d'un système d'inégalités à satisfaire. Si nous souhaitons que ce système soit linéaire, il faut choisir un ensemble d'incertitude polyédrique et une norme appropriée. Cette norme doit pouvoir être décrite par un nombre fini d'inégalités linéaires, comme la norme ℓ_∞ ou ℓ_1 , ou une combinaison des deux, voire une norme induite par un borné avec un intérieur non vide. Les choix de l'ensemble d'incertitude et de la norme sont a priori indépendants. Il y a cependant une certaine logique à lier les deux et plus particulièrement à utiliser la norme induite par l'ensemble d'incertitude lui-même.

Pour illustrer le propos de manière concrète, nous considérons le cas d'un ensemble d'incertitude défini comme l'intersection de la boule de rayon k dans la norme ℓ_∞ et de la boule de rayon $k\sqrt{p}$ dans la norme ℓ_1 . Rendons explicite la dépendance de cet ensemble par rapport à k en écrivant

$$\Xi(k) = \{\xi \in \Omega \mid \|\xi\|_1 \leq k\sqrt{p}, \|\xi\|_\infty \leq k\}.$$

Le polyèdre $\Xi(k)$ induit la norme

$$\delta(\mu) = \min\{k \mid \mu \in \Xi(k)\} = \max\left\{\frac{1}{\sqrt{p}}\|\mu\|_1, \|\mu\|_\infty\right\}.$$

Pour trouver la distance d'un point ξ à l'ensemble Ξ il faut résoudre le problème d'optimisation

$$d(\xi, \Xi) = \min_{\nu} \{\delta(\xi - \nu) \mid \nu \in \Xi\}.$$

3.9. EXTENSION DU CONCEPT DE ROBUSTESSE : VIOLATION CONTRÔLÉE DES CONTRAINTES

La figure 3.4 illustre la distance de ξ à l'ensemble Ξ dans la norme composite choisie. Les courbes en traits tiretés représentent des ensembles de points équidistants de ξ . Le point $\nu \in \Xi$ est un des points les plus proches de ξ . (Ce point n'est pas unique parce que la fonction distance choisie n'est pas strictement convexe.) La distance calculée est $d(\xi, \Xi) = \delta(\nu, \xi) = 2.53$.

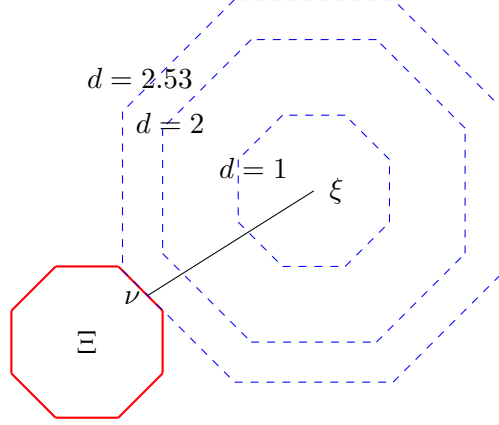


FIG. 3.4 – Distance $d(\xi, \Xi)$ dans la norme composite

La question posée est de savoir si le problème de maximisation qui intervient dans la définition de l'équivalent robuste peut se formuler comme un programme linéaire. Pour répondre affirmativement, nous opérons quelques transformations. Commençons par la plus simple qui introduit une variable auxiliaire ζ . Cela donne

$$\begin{aligned} \max_{\xi, \nu, \zeta} \quad & (P^T x)^T \xi - \alpha \zeta \\ & \delta(\xi - \nu) \leq \zeta \\ & \nu \in \Xi, \xi \in \Omega. \end{aligned}$$

La maximisation de l'objectif force ζ à prendre une valeur aussi petite que possible. Cela assure l'égalité $\delta(\mu) = \zeta$.

La transformation suivante vise à exprimer que ν appartient à Ξ . Pour cela, il faut faire intervenir les valeurs absolues des composantes, ce que nous réalisons classiquement par une décomposition en parties positive et négative. De même, la norme δ que nous avons choisie peut s'exprimer à l'aide des valeurs absolues des composantes de $\mu = \xi - \nu$. Ces décompositions $\mu = \mu^+ - \mu^-$ et $\nu = \nu^+ - \nu^-$ vont nous permettre d'éliminer μ , ν et ξ de la formulation. Nous pourrions alors écrire le test de robustesse comme

$$\begin{aligned} \max_{\mu^+, \mu^-, \nu^+, \nu^-} \quad & (P^T x)^T [(\mu^+ - \mu^-) + (\nu^+ - \nu^-)] - \alpha \zeta \\ & 0 \leq \mu^+ \leq \zeta \\ & 0 \leq \mu^- \leq \zeta \\ & (\mu^+ + \mu^-)^T e \leq \zeta \sqrt{p} \\ & 0 \leq \nu^+ \leq k \\ & 0 \leq \nu^- \leq k \\ & (\nu^+ + \nu^-)^T e \leq k \sqrt{p}. \end{aligned} \tag{3.18}$$

3.9. EXTENSION DU CONCEPT DE ROBUSTESSE : VIOLATION CONTRÔLÉE DES CONTRAINTES

Pour qu'une solution optimale μ^+ et μ^- de (3.18) corresponde bien à la décomposition de μ en parties positive et négative, il faut que la condition de complémentarité $\mu_i^+ \mu_i^- = 0$ soit vérifiée pour chaque coordonnée $i = 1, \dots, p$. (La même condition doit être vérifiée pour ν .) Il est facile de voir que si ce n'était pas le cas à l'optimum, il serait toujours possible de définir la décomposition alternative $\mu = \tilde{\mu}^+ - \tilde{\mu}^-$, avec $\tilde{\mu}_i^+ = \mu_i^+ - \min\{\mu_i^+, \mu_i^-\} \geq 0$ et $\tilde{\mu}_i^- = \mu_i^- - \min\{\mu_i^+, \mu_i^-\} \geq 0$. Cette décomposition satisfait la condition $\mu_i^+ \mu_i^- = 0$ tout en préservant l'ensemble des autres contraintes; quant à l'objectif, sa valeur est inchangée. Le même raisonnement peut s'appliquer à ν^+ et ν^- . Il existe donc une solution optimale pour laquelle nous avons bien la décomposition en parties positive et négative. L'équivalence est donc établie.

Le dual du problème linéaire (3.18) est

$$\begin{aligned} \min_{u, v, \rho, \sigma} \quad & k(v^+ + v^-) + \sigma k\sqrt{p} \\ & (u^+ + u^-)^T e + \rho\sqrt{p} \leq \alpha \\ & \rho e + u^+ \geq P^T x \\ & \rho e + u^- \geq -P^T x \\ & \sigma e + v^+ \geq -P^T x \\ & \sigma e + v^- \geq P^T x \\ & u^+, u^-, v^+, v^-, \rho, \sigma \geq 0. \end{aligned}$$

L'équivalent robuste étendu de la contrainte $a(x)^T \xi \leq 0$ est donc l'ensemble d'inégalités sur les variables x, u, v, σ

$$\begin{aligned} k(v^+ + v^-) + \sigma k\sqrt{p} & \leq -\bar{a}^T x \\ (u^+ + u^-)^T e + \rho\sqrt{p} & \leq \alpha \\ \rho e + u^+ & \geq P^T x \\ \rho e + u^- & \geq -P^T x \\ \sigma e + v^+ & \geq -P^T x \\ \sigma e + v^- & \geq P^T x \\ u^+, u^-, v^+, v^-, \rho, \sigma & \geq 0. \end{aligned} \tag{3.19}$$

La robustesse étendue fait appel à un paramètre α de tolérance de la violation. Plus ce paramètre est petit, plus la violation de la contrainte est restreinte. Si α tend vers l'infini, nous retrouvons la solution robuste standard pour laquelle la contrainte n'est plus prise en considération si l'aléa se trouve hors de l'ensemble d'incertitude.

Dans le test de robustesse (3.18), l'ensemble Ω est tout \mathbb{R}^p . En pratique Ω est un ensemble plus restreint. Par exemple, si le paramètre incertain est une quantité physique, cette quantité peut ne prendre que des valeurs positives. Cela induit une contrainte sur ξ . Prenons le cas simple d'un ensemble $\Omega = \{\xi \mid \underline{\xi} \leq \xi \leq \bar{\xi}\}$ défini par des contraintes de boîte. Le test de robustesse étendue est obtenu en résolvant le problème (3.18) auquel nous ajoutons la contrainte

$$\underline{\xi} \leq (\mu^+ - \mu^-) + (\nu^+ - \nu^-) \leq \bar{\xi}.$$

Cette double contrainte ajouterait deux variables supplémentaires (w^+, w^-) dans le dual. Les

contraintes garantissant une solution robuste étendue deviennent alors

$$\begin{aligned}
 k(v^+ + v^-) + \sigma k\sqrt{p} + (\bar{\xi} - \underline{\xi})^T(w^+ + w^-) &\leq -\bar{a}^T x \\
 (u^+ + u^-)^T e + \rho\sqrt{p} &\leq \alpha \\
 \rho e + u^+ + w^+ &\geq P^T x \\
 \rho e + u^- - w^- &\geq -P^T x \\
 \sigma e + v^+ + w^+ &\geq -P^T x \\
 \sigma e + v^- - w^- &\geq P^T x \\
 u^+, u^-, v^+, v^-, w^+, w^-, \rho, \sigma &\geq 0.
 \end{aligned}$$

Nous retrouvons les mêmes contraintes que dans (3.19), mais avec des variables supplémentaires. L'ensemble des solutions admissibles est donc plus grand, ce qui est conforme au fait que le contrôle de la violation de contraintes se fait sur un ensemble plus restreint d'aléas.

Un dernier mot concernant la dimension de la contre-partie robuste étendue. Pour une contrainte déterministe nous avons introduit 6 variables $(u^+, u^-, v^+, v^-, w^+, w^-)$ dans \mathbb{R}_+^p et deux variables scalaires (ρ, σ) , soit $6p + 2$ variables. Il y a 4 groupes de contraintes supplémentaires, chaque membre du groupe représentant p contraintes individuelles, et il y a aussi une contrainte isolée, soit au total $4p + 1$. Cet accroissement est significatif, mais *il demeure linéaire dans la dimension de l'aléa ξ* .

3.10 Conclusion

Dans ce chapitre, nous avons décrit les bases de l'optimisation robuste dans un cadre linéaire. Nous avons discuté deux extensions importantes, l'une concernant les systèmes dynamiques, l'autre une façon de prendre en compte les cas exceptionnels où l'aléa n'appartient pas à l'ensemble d'incertitude.

D'autres extensions importantes ont été développées dans la littérature. (Voir en particulier [11].) Elles concernent par exemple le cas d'ensembles d'incertitude non linéaires et de problèmes dont la version déterministe est aussi non linéaire. Malgré la complexité introduite par la non-linéarité, il est possible d'obtenir des contreparties robustes polynomiales en faisant des hypothèses sur ces non-linéarités. Essentiellement, nous demandons que ces non-linéarités soient décrites par des cônes : le cône du second-ordre ou celui des matrices positives semi-définies. Le résultat est en gros que la contrepartie robuste est un problème conique que nous pouvons résoudre par une méthode polynomiale de points intérieurs.

Il est intéressant de noter que la dérivation de la contrepartie robuste d'une contrainte repose entièrement sur le fait que cette contrainte est une fonction linéaire (affine) des paramètres incertains. Si les coefficients de ces paramètres sont eux-mêmes des fonctions linéaires de la variable de décision x , la contrepartie robuste est un ensemble de contraintes linéaires en x et dans des variables auxiliaires. Cela dit, rien n'empêche d'ajouter des contraintes supplémentaires sur x , comme d'être un vecteur de nombres entiers. Nous pouvons ainsi inclure dans l'optimisation robuste des problèmes en nombres entiers. Cette remarque s'applique sans restriction aux problèmes statiques. Dans le cas dynamique, il avait fallu avoir recours à des règles de décision

linéaires dans les paramètres. Cela veut dire que les variables de décision de recours, c'est-à-dire prises aux étapes 2 et suivantes, ne peuvent pas être restreintes à des valeurs entières dans cette formulation.

Bibliographie

- [1] E. Adida and G. Perakis. A Robust Optimization Approach to Dynamic Pricing and Inventory Control with no Backorders. *Mathematical Programming*, 107(1-2) :97 – 129, 2006.
- [2] R. Apparigliato, J.-P. Vial, and R. Zorgati. Optimisation Robuste Linéaire : Application à la Gestion Court Terme d'une Vallée Hydraulique. Working paper H-R32-2007-00658-FR, EDF R&D, Département Osiris, 1, avenue du Général de Gaulle, F-92141 Clamart, 2007.
- [3] A. Ben-Tal, D. Bertsimas, L. El Ghaoui, A. Nemirovski, and M. Sim. *Robust Optimization*, volume Book in progress. 2007.
- [4] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending the Scope of Robust Optimization : Comprehensive Robust Counterparts of Uncertain Problems. *Mathematical Programming Series B*, 107(1-2) :63–89, 2006.
- [5] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-Ph. Vial. Retailer-Supplier Flexible Commitments Contracts : a Robust Optimization Approach. *Manufacturing and Service Operations Management*, 7 :248–271, 2005.
- [6] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming*, 99(2) :351–376, 2004.
- [7] A. Ben-Tal and A. Nemirovski. Robust Convex Optimization. *Mathematics of Operations Research*, 23 :769 – 805, 1998.
- [8] A. Ben-Tal and A. Nemirovski. Robust Solutions of Linear Programming Problems Contaminated with Uncertain Data. *Mathematical Programming Series A*, 88 :411 – 424, 2000.
- [9] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization : Analysis, Algorithms and Engineering Applications*. MPS-SIAM series on Optimization. Society for Industrial and Applied Mathematics and Mathematical Programming Society, 2001.
- [10] A. Ben-Tal and A. Nemirovski. On Polyhedral Approximations of the Second-Order Cone. *Mathematics of Operations Research*, 26(2) :193–205, 2001.
- [11] A. Ben-Tal and A. Nemirovski. Selected Topics in Robust Convex Optimization. Technical report, Faculty of Industrial Engineering and Management, Technion, Technion city, Haifa 32000, Israel, 2006. (To appear in *Mathematical Programming*).
- [12] D. Bertsimas, D. Pachamanova, and M. Sim. Robust Linear Optimization under General Norms. *Operations Research Letters*, 32 :510–516, 2004.
- [13] D. Bertsimas and M. Sim. Price of Robustness. *Operations Research*, 52 :35–53., 2004.
- [14] D. Bertsimas and A. Thiele. A Robust Optimization Approach to Inventory Theory. *Operations Research*, 54(1) :150–168, 2006.
- [15] Daniel Bienstock. Experiments in Robust Portfolio Optimization. Technical report, Center for Financial Engineering, Columbia University, January 2007.

- [16] P. Bonami and M. A. Lejeune. An Exact Solution Approach for Portfolio Optimization Problems under Stochastic and Integer Constraints. Technical report, 2007.
- [17] G. Calafiore and M.C. Campi. Decision Making in an Uncertain Environment : a Scenario based Optimization Approach. In J. Andrysek, M. Karny, and J. Kracik, editors, *Multiple Participants Decision Making*, pages 99–111. Advanced Knowledge International, 2004.
- [18] G. C. Calafiore. Ambiguous Risk Measures and Optimal Robust Portfolios. Technical report (to appear in *SIAM Journal on Optimization* 2007), Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy, 2007.
- [19] G. C. Calafiore and L. El-Gahoui. On Distributionally Robust Chance-Constrained Linear Programs. *Journal of Optimization Theory and Applications*, 130 :1–22, 2006.
- [20] S. Ceria and R. Stubbs. Incorporating Estimation Errors into Portfolio Selection : Robust Portfolio Construction. *Journal of Asset Management*, 7(2) :109 – 127, 2006.
- [21] A. Charnes and W.W. Cooper. Chance Constrained Programming. *Management Science*, 6 :73–89, 1959.
- [22] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost Horizons and Certainty Equivalents : an Approach to Stochastic Programming of Heating Oil. *Management Science*, 4 :235–263, 1958.
- [23] X. Chen, M. Sim, and P. Sun. A Robust Optimization Perspective of Stochastic Programming. Working paper, NUS Business school, 2005.
- [24] A. Prékopa. *Stochastic Programming*. Kluwer, Dordrecht, Boston, 1995.
- [25] L. El-Ghaoui and H. Lebret. Robust Solutions to Least-square Problems to Uncertain Data Matrices. *SIAM Journal of Matrix Analysis and Applications*, 18 :1035–1064, 1997.
- [26] S. J. Garstka and R. J.-B. Wets. On Decision Rules in Stochastic Programming. *Mathematical Programming*, 7 :117–143, 1974.
- [27] D. Goldfarb and G. Iyengar. Robust Portfolio Selection Problems. *Mathematics of Operations Research*, 28 :1 – 38, 2003.
- [28] R. Henrion. Structural Properties of Linear Probabilistic Constraints. *A Journal of Mathematical Programming and Operations Research*, 56 :425 – 440, 2007.
- [29] C.C. Holt, F. Modigliani, and H.A. Simon. A Linear Decision Rule for Production and Employment Scheduling. *Management Science*, 2(1) :1–30, 1955.
- [30] G. Iyengar. Robust Dynamic Programming. *Mathematics of Operations Research*, 30 :257 – 280, 2005.
- [31] A. Nemirovski and A. Shapiro. Scenario Approximations of Chance Constraints. Preprint, School of Industrial and Systems Engineering, Georgia Institute of Technology, 2004.
- [32] A. Nemirovski and A. Shapiro. Convex Approximations of Chance Constrained Programs. *SIAM Journal on Optimization*, 17 :969–996, 2006.
- [33] Yu. Nesterov and A. Nemirovski. *Interior-Point Polynomial Methods in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.
- [34] A. Ouorou and J.-P. Vial. A Model for Robust Capacity Planning for Telecommunications Networks under Demand Uncertainty. In *submitted to the 6th International Workshop on Design and Reliable Communication Networks, DRCN 2007*, 2007.

BIBLIOGRAPHIE

- [35] J. Pinter. Deterministic Approximations of Probability Inequalities. *ZOR Methods and Models of Operations Research, Series Theory*, 33 :219–239, 1989.
- [36] A. Prékopa. Probabilistic Programming. In A. Ruszczyński and A. Shapiro, editors, *Handbooks in Operations Research and Management Science*, volume 10. North-Holland, 2003.
- [37] A. Shapiro and A. Nemirovski. On Complexity of Stochastic Programming Problems. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*. Springer, London, 2005.
- [38] A. L. Soyster. Convex Programming with Set-inclusive Constraints and Applications to Inexact Linear Programming. *Operations Research*, 21 :1154–1157, 1973.
- [39] J.F. Sturm. Using SeDuMi 1.02, a Matlab Toolbox for Optimization over Symmetric Cones (Updated for Version 1.05). Technical report, 2001.
- [40] J.-P. Vial and R. Apparigliato. Optimisation Robuste Linéaire - Une Introduction Succinte. Technical report, EDF R&D, Dept OSIRIS, H-R32-2006-04801-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [41] J.-P. Vial, R. Apparigliato, and F. Babonneau. Introduction to Linear Robust Optimization. In F. S. Hillier, editor, *Handbook on Uncertainty and Environmental Decision Making*, International Series in Operations Research and Management Science. Springer, to appear (submitted).

Chapitre 4

L'optimisation robuste appliquée à la gestion court-terme d'une vallée hydraulique

Sommaire

4.1	Introduction	112
4.2	Modèle déterministe de gestion d'une vallée hydraulique	114
4.2.1	Modélisation choisie	114
4.2.2	Modèle déterministe	116
4.3	Incertitude	117
4.3.1	Incertitude sur les apports d'eau	117
4.3.2	Les recours	119
4.3.3	Optimisation, incertitude et validation	119
4.4	Politique de gestion Robuste à Révisions Périodiques (RRP) du problème hydraulique	121
4.5	Politique de gestion Robuste à Ajustements Linéaires (RAL) du problème hydraulique	122
4.5.1	Contraintes de turbinage	124
4.5.2	Contraintes de volume	125
4.5.3	Fonction-objectif	127
4.6	Tailles des modèles	127
4.7	Etude expérimentale	130
4.7.1	Implémentation	130
4.7.2	Jeu d'essais	131
4.7.3	Simulation	132
4.8	Expérimentations numériques	135
4.8.1	Comparaison entre DF et DRP	135
4.8.2	Comparaison entre DRP et RRP	136
4.8.3	Comparaison entre RRP et RAL	140
4.9	Expérimentations numériques : débit minimum	144
4.10	Conclusion	145
	Bibliographie	146

Ce chapitre est une application de l'optimisation robuste au problème de gestion d'une vallée hydraulique à l'horizon hebdomadaire. Les concepts du chapitre précédent ont été appliqués sur un jeu d'essais type de 3 réservoirs et 6 groupes sur 7 jours. Les résultats montrent qu'il est possible de réduire de façon conséquente les violations de contrainte de volume sans un surcoût excessif. Une maquette (Matlab, GLPK) a été implémentée et permet de tester toutes les approches proposées sur diverses vallées hydrauliques. Cette contribution a donné lieu à une note interne EDF [1], une proposition de chapitre dans [7] et à de nombreuses présentations orales lors de conférences ou séminaires [6, 4, 2, 5, 3].

4.1 Introduction

La gestion des stocks hydrauliques est un enjeu capital pour les producteurs d'électricité. En effet, les réserves hydrauliques peuvent être vues comme les uniques stocks en électricité (hors contrats), disponibles en peu de temps, et dont le coût proportionnel est nul. Nous sommes donc face à une ressource capitale qu'il est nécessaire de gérer au mieux. De par ses propriétés, l'hydraulique est utilisée en substitution aux unités de production les plus chères. Mais pour que cette gestion soit réellement optimisée, il est nécessaire d'avoir une planification sur un horizon suffisamment important (annuel) pour prévoir les différentes périodes de pointe (forte demande en électricité en hiver (chauffage) et en été (climatisation)) et anticiper la variabilité des apports naturels. En effet, une grande quantité d'apports peut obliger les usines à turbiner afin d'éviter des déversements ou pour assurer un débit minimum à la vallée, tout comme un manque d'apports peut inciter un arrêt de la production .

L'approche utilisée pour résoudre ce problème d'optimisation sous aléa dépend de multiples paramètres : la taille du parc hydraulique, la connaissance ou non de la distribution des apports, le niveau de modélisation des unités de production,... Dans les pays à forte capacité hydraulique (Norvège, Brésil,...), l'approche la plus utilisée est la programmation dynamique stochastique duale [17, 15, 19]. Cette approche leur permet de résoudre un problème avec un nombre de stocks supérieur à ce que la programmation dynamique stochastique standard leur permettrait de résoudre (autour de la dizaine sur 30 à 60 pas de temps). D'autre part, les stocks hydrauliques étant de grande taille, les unités de type thermique sont peu nombreuses ce qui permet de résoudre le problème frontalement. Unger [21] propose de résoudre le problème de gestion d'un réservoir dans un contexte de couverture des risques et de gestion d'un portefeuille et de production. Les décisions de turbinage/pompage, assimilées à des options (swing), sont des fonctions (en escalier) dépendant du prix d'exercice, de la demande,... Une telle approche, couplant directement marchés financiers et hydraulique n'est pas transposable à EDF qui a encore une vision centralisée. Une approche par programmation stochastique où l'aléa d'hydraulicité est représenté par un arbre a été étudiée par Römissh et al [16, 14, 13, 12, 10] ou Fleten et Kristoffersen [11]. Le problème de la programmation stochastique est qu'une représentation détaillée de l'aléa nécessite un grand nombre de branches, ce qui combiné avec une croissance exponentielle dans le temps, conduit à une taille d'arbre très importante. Or, l'horizon hebdomadaire comportant 7 à 12 jours à pas de temps bi-horaire, une telle approche n'est pas numériquement traitable.

Prékopa a introduit une approche basée sur l'utilisation de contraintes en probabilité [18]. Cette approche a pour but de satisfaire les contraintes de bornes sur les volumes avec une certaine probabilité. En théorie, cette approche correspond à nos besoins. Mais, pour transformer les contraintes en probabilité en contraintes linéaires, il est nécessaire de calculer des quantiles de la distribution des apports. Il est donc nécessaire de connaître la distribution de probabilité des apports hydrauliques. En pratique, cette distribution est inconnue ce qui rend cette approche difficile à mettre en place dans notre cas. On peut rajouter que même si cette distribution était connue, la convexité (et la connexité) de ces contraintes est loin d'être assurée... Une façon d'approximer cette contrainte en probabilité par des contraintes convexes tout en proposant une approche opérationnelle est d'appliquer l'optimisation robuste, développée initialement pour des problématiques d'approvisionnement par Ben-Tal et Nemirovski [8]. Le succès de cette approche dans de très nombreux domaines (cf. chapitre 3) nous conforte dans ce choix.

L'optimisation robuste propose une couverture contre les aléas qui semblent les plus réalistes, excluant donc le cas des événements extrêmes. Cet ensemble d'événements les plus réalistes est appelé ensemble d'incertitude. L'optimisation robuste nous assure donc que, pour toute réalisation de l'aléa dans cet ensemble d'incertitude, la solution obtenue sera réalisable. Toute la difficulté réside dans la construction de cet ensemble puisque la nature et la forme de l'ensemble doivent être cohérents avec la réalité. Les aléas d'hydraulicité étant supposés indépendants d'une période à l'autre et d'un réservoir à l'autre, nous avons choisi dans cette étude de considérer comme ensemble d'incertitude une boule ℓ_2 , qui est approximée par une intersection de normes ℓ_1 et ℓ_∞ . Avec un tel ensemble polyédrique, la version robuste du problème est encore un programme linéaire, de taille légèrement plus grande, et donc de complexité voisine de celle du modèle déterministe. Cette modélisation a cependant l'inconvénient majeur de ne pas prendre en compte les possibilités d'ajustement (ou de recours) attachées aux périodes futures. Pour remédier à cette limitation, les auteurs de [9] ont utilisé le concept de règle de décision linéaire, ce qui conduit de nouveau à la formulation du problème stochastique global comme un programme linéaire, certes de plus grande taille, mais encore gérable par les outils de modélisation et de résolution actuellement disponibles. Dans le cas de l'hydraulique, les règles de décision prennent la forme de fonctions linéaires des observations journalières sur les apports d'eau.

Cette approche robuste a été testée sur une vallée de 3 réservoirs et 6 groupes, gérée sur une période de 7 jours (84 pas de temps). Cette vallée, quoique de petite taille, demeure toutefois représentative d'une partie des vallées hydrauliques françaises. Il n'existe pas de solution reconnue optimale à laquelle nous puissions comparer la solution robuste. Pour valider notre approche, nous avons procédé à des simulations. Cette phase de validation est généralement peu mise en avant dans la littérature. Cependant, elle semble obligatoire si nous voulons évaluer la qualité de la solution. Pour cela, nous avons tiré 10 000 scénarios et avons mis en œuvre deux approches différentes :

- Dans l'approche *robuste à révisions périodiques* (RRP), seules les contraintes de la première journée sont rendues robustes. Pour les journées suivantes, les contraintes ne s'appliquent que pour les valeurs moyennes des apports d'eau. La révision périodique consiste à lancer une optimisation au début de chaque journée, en tenant compte des niveaux de volume actualisés. La révision périodique introduit une capacité d'adaptation. La prise en compte dans l'optimisation d'un horizon dépassant le premier jour—même sur un mode déterministe— donne une cohérence aux décisions du premier jour.

- Dans l'approche *robuste à ajustements linéaires* (RAL), des règles de décision pour les jours 2 et suivants introduisent directement dans le problème d'optimisation la faculté d'ajustement aux aléas manifestés. Le modèle résultant est plus complexe, mais il est théoriquement implémentable sur tout l'horizon d'étude sans réactualisation.

Comme élément de comparaison, nous avons retenu tout d'abord une politique à information parfaite (IP). Dans cette politique, l'optimisation se fait sur chaque scénario avec connaissance intégrale de la chronique des apports d'eau. Cette politique est une référence idéale, non réalisable. Elle n'entraîne aucune violation de contrainte et sa performance économique est une borne inférieure absolue pour toute autre politique. Le deuxième élément de comparaison est une politique que nous nommons *déterministe à révision périodique* (DRP). Cette politique peut être considérée comme assez proche de celle suivie en pratique en exploitation : à chaque période, nous optimisons sur une version déterministe du problème, où les apports d'eau sont fixés à leur valeur moyenne. Cette politique possède une certaine flexibilité qui la rend un peu plus réaliste¹ qu'une *politique déterministe fixe* (DF).

Les expériences numériques montrent que les approches robustes assurent une excellente satisfaction des contraintes. Cela est particulièrement vrai pour RRP, mais, de manière surprenante, aussi pour RAL où toutes les décisions sont fixées (sous forme de règles) à l'origine de la période de planification. Le surcoût par rapport à IP n'est que de 0,57 % et 1,15 %, pour respectivement RRP et RAL. Par contraste, DRP est incapable de satisfaire les contraintes. En particulier, nous observons des déversés sur presque tous les scénarios et la valeur moyenne de ces déversés est si importante que la performance économique sur les simulations ne peut être prise en considération².

4.2 Modèle déterministe de gestion d'une vallée hydraulique

4.2.1 Modélisation choisie

Le parc hydraulique est organisé en une quarantaine de vallées nationales, chaque vallée étant modélisée comme un ensemble cohérent de réserves hydrauliques et d'usines de production, interconnectées (Figure 4.1). Il représente plus de 100 usines hydrauliques. Chaque vallée est caractérisée par sa topologie qui décrit les interconnexions entre réserves et usines, par la nature des usines (pompage ou turbinage) et par certains paramètres comme le délai de parcours de l'eau entre deux réservoirs. En pratique, une vallée est sujette à de nombreuses contraintes :

- contraintes de bornes sur les turbines/pompes de chaque groupe ;
- contraintes de bornes sur les volumes des réservoirs ;
- paliers de production ;
- durée minimale de palier de production et gradient de débit ;

La gestion d'une vallée doit tenir compte de ces contraintes, des équations de flot qui régissent la dynamique du système et enfin du coût marginal de l'électricité et de la valeur finale d'usage de

¹Une politique déterministe fixe (DF) où les décisions de période 2 et suivantes seraient fixées rigidement dès la première période, donc sans procéder à des réoptimisations ultérieures, est totalement inopérante, et viole massivement les contraintes.

²En effet, nous observons que les violations importantes des contraintes permettent d'améliorer le coût de IP !

4.2. MODÈLE DÉTERMINISTE DE GESTION D'UNE VALLÉE HYDRAULIQUE

l'eau. Dans la pratique, l'indisponibilité des centrales, planifiée ou subie, est un facteur important intervenant dans la gestion.

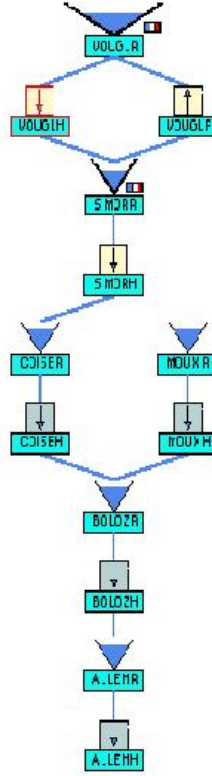


FIG. 4.1 – Représentation de la vallée de l'Ain

Dans notre cas, pour alléger les écritures, nous ne considérerons plus la notion d'usine et tous les groupes seront gérés comme s'ils étaient autonomes. De plus, pour modéliser le problème de façon linéaire et sans caractère discret, nous ne considérerons que certaines contraintes du problème (tout comme dans le modèle APOGEE [22, 23, 24]. Les contraintes prises en compte sont les contraintes de borne sur les volumes et sur les commandes et la contrainte d'évolution des réservoirs. Enfin, nous supposerons que toutes les centrales sont toujours disponibles et que seul le placement en puissance est traité (pas de télé réglage, de réserve primaire ou tertiaire).

D'autre part, nous faisons abstraction dans cette étude des aléas sur les prix, ce qui se traduit par une chronique de coûts marginaux fixes. Le seul aléa du problème est l'hydraulicité, c'est à dire la quantité d'apports d'eau qui arrive à chaque pas de temps et chaque réservoir par des causes naturelles comme la pluie, la fonte des neiges, etc. L'hypothèse de coûts déterministes est forte mais, outre le fait qu'elle est actuellement pratiquée en opération sur le court terme, elle nous permet d'approcher de manière intéressante les aléas sur les apports d'eau.

4.2.2 Modèle déterministe

La version déterministe du problème correspond au cas où la chronique des apports est connue et certaine. Nous supposons connaître une chronique de prévision d'apports hydrauliques par réservoir.

Paramètres et variables

- H : horizon de planification
- r : réservoir r
- R : ensemble des réservoirs
- g : groupe (turbinage ou pompage) g
- $o(g)$: réservoir amont du groupe
- $e(g)$: réservoir aval du groupe
- V_r^0 : volume d'eau initial dans le réservoir r
- a_r^t : apport d'eau prévu au réservoir r à la date t
- \underline{V}_r^t : volume minimal autorisé au réservoir r à la date t
- \bar{V}_r^t : volume maximal autorisé au réservoir r à la date t
- \underline{F}_g^t : flot minimal autorisé dans le groupe g à la date t
- \bar{F}_g^t : flot maximal autorisé dans le groupe g à la date t
- ρ_g : taux de transformation du flot en électricité
- c_t : coût marginal de l'électricité (donné par le coordinateur)
- ω_r^0 : valeur de l'eau à l'état initial
- ω_r^H : valeur de l'eau à l'horizon

Il y a une seule classe de variables dans le problème :

$$F_{gt} : \text{flot dans le groupe } g \text{ durant la période } t.$$

L'altitude induit un ordre partiel sur les réservoirs. Nous supposons que l'ensemble R respecte cet ordre. Cela permet de repérer facilement si le groupe g effectue du pompage ou du turbinage. Nous affectons d'un signe le taux de transformation des flots :

$$\rho_g = \begin{cases} < 0 & \text{si } o(g) < e(g) \Rightarrow \text{groupe de turbinage (produit de l'électricité)} \\ > 0 & \text{si } o(g) > e(g) \Rightarrow \text{groupe de pompage (consomme de l'électricité)}. \end{cases}$$

Les paramètres suivants facilitent l'écriture des contraintes. Tout d'abord la somme cumulée des apports

$$A_r^t = \sum_{\tau \leq t} a_r^\tau.$$

L'autre paramètre est un indicateur qui reflète le type d'association d'un groupe à un réservoir.

$$\delta_{gr} = \begin{cases} \text{sign}(\rho_g) & \text{si } g \text{ est connecté à } r \\ 0 & \text{autrement.} \end{cases}$$

Le problème comporte des variables de commandes, les flots turbinés F_{gt} , et des variables d'état, les volumes dans les réservoirs, V_r^t .

Formulation déterministe

Le volume du réservoir à la fin de la période t est donné par :

$$V_r^t = V_r^0 + \sum_{\tau \leq t} \left(a_r^\tau + \sum_g \delta_{gr} F_{g\tau} \right) = V_r^0 + A_r^t + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau}.$$

L'impact économique du groupe g est donné par le coût :

$$\rho_g \sum_{\tau \leq H} c_\tau F_{g\tau}.$$

Nous rappelons que $\rho_g < 0$ (coefficient de rendement du groupe) pour du turbinage. Le turbinage décroît les coûts.

L'impact économique de la variation de volume d'eau dans les réservoirs est donné par le coût

$$\sum_r \omega_r^0 V_r^0 - \omega_r^H V_r^H = \sum_r (\omega_r^0 - \omega_r^H) V_r^0 + \sum_r \omega_r^H (V_r^0 - V_r^H).$$

L'optimiseur arbitre donc à chaque instant entre la conservation de l'eau rémunérée à la valeur de l'eau et l'utilisation de l'eau rémunérée au coût marginal du système (c-à-d le coût de l'unité de production la moins chère à utiliser dans le cas d'une unité de demande supplémentaire).

Par la suite, nous omettrons la contribution fixe $\sum_r (\omega_r^0 - \omega_r^H) V_r^0$ et nous utiliserons la notation $\omega_r = \omega_r^H$.

Nous pouvons donc formuler le problème

$$\min \quad z \tag{4.1a}$$

$$z \geq \sum_g \rho_g \sum_{\tau \leq T} c_\tau F_{g\tau} + \sum_r \omega_r (V_r^0 - V_r^H) \tag{4.1b}$$

$$V_r^t = V_r^0 + A_r^t + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau}, \quad \forall r, \quad \forall t \leq H \tag{4.1c}$$

$$\underline{V}_r^t \leq V_r^t \leq \bar{V}_r^t, \quad \forall r, \quad \forall t \leq H \tag{4.1d}$$

$$\underline{F}_g^t \leq F_{gt} \leq \bar{F}_g^t, \quad \forall g, \quad \forall t \leq H. \tag{4.1e}$$

4.3 Incertitude

4.3.1 Incertitude sur les apports d'eau

Nous souhaitons prendre en compte les incertitudes sur les apports hydrauliques dans le problème déterministe. L'apport d'eau au réservoir r et au pas de temps t est une variable aléatoire que nous pouvons écrire

$$\tilde{a}_r^t = a_r^t + \sigma_r^t \epsilon_r^t.$$

Nous interprétons a_r^t comme la valeur moyenne, et ϵ_r^t comme le facteur de déviation autour de la moyenne, exprimé comme le multiple de l'écart-type σ_r^t . L'aléa du problème est désormais le bruit blanc ϵ_r^t qui est centré et réduit.

Hypothèse 4.1 *Nous supposons que tous les aléas ϵ_r^t sont indépendants entre eux.*

Ce modèle simplifié de l'incertitude a été retenu pour être compatible avec les générateurs de simulation d'apports existants. Nous pouvons cependant envisager des modèles plus complets (comme des modèles autorégressifs) sans rendre la modélisation du problème d'optimisation robuste plus complexe et sans augmenter de manière incontrôlée la taille des problèmes linéaires à résoudre.

Par ailleurs, les apports d'eau sont reçus par pas de temps et par réservoir, mais ne sont pas directement observables par le décideur. Nous effectuons donc l'hypothèse suivante :

Hypothèse 4.2 *La seule information sur les apports d'eau qui soit accessible au planificateur est celle des apports cumulés sur tous les pas de temps d'une journée pour chaque réservoir.*

Suivant cette hypothèse, pour un modèle avec 12 pas de temps par jour, l'information disponible durant les pas de temps 25 à 36 est la même en chaque pas de temps et correspond au cumul des apports des jours 1 (pas de temps 1 à 12) et 2 (pas de temps de 13 à 24). Il est commode d'introduire une notation spécifique pour repérer le jour où est révélée la plus récente information sur les apports d'eau. Ce jour se calcule simplement par la formule :

$$J(t) = \left\lceil \frac{t}{T} \right\rceil - 1,$$

où $\lceil a \rceil$ désigne la partie entière supérieure de a .

Les apports d'eau cumulés dans le réservoir r durant le jour j s'écrivent :

$$\begin{aligned} \tilde{A}_r^j &= \sum_{t=(j-1)T+1}^{jT} (a_r^t + \sigma_r \epsilon_r^t) \\ &= \left(\sum_{t=(j-1)T+1}^{jT} a_r^t \right) + \sigma_r \left(\sum_{t=(j-1)T+1}^{jT} \epsilon_r^t \right). \end{aligned} \quad (4.2)$$

Nous pouvons récrire cette variable aléatoire \tilde{A}_r^j comme une perturbation autour d'une valeur nominale (tout comme pour ϵ_r^t) :

$$\tilde{A}_r^j = \left(\sum_{t=(j-1)T+1}^{jT} a_r^t \right) + \sigma_r \sqrt{T} \theta_r^j, \quad (4.3)$$

où nous pouvons considérer que tous les $\theta_r^j = \sum_{t=(j-1)T+1}^{jT} \epsilon_r^t$ sont, comme les ϵ_r^t , des bruits blancs centrés réduits et indépendants entre eux.

Pour résumer, nous distinguons dans notre modèle deux aléas qui s'avèrent être des bruits blancs centrés, réduits et indépendants entre eux :

- ϵ_r^t qui représente la réalisation de l'aléa en chaque instant ;
- θ_r^j qui représente l'observation de l'aléa en fin de chaque journée.

Remarque 4.1 *En pratique, le modèle d'apports utilisé est un modèle qui génère des perturbations autour d'une moyenne à l'aide d'une loi normale bi-tronquée. Le fait de connaître cette information ainsi que les paramètres de cette loi (écart-type, moyenne) pourra être exploité pour affiner la prise en compte de l'incertitude dans le processus de décision par optimisation robuste.*

4.3.2 Les recours

Le problème de détermination des commandes hydrauliques à l'instant initial pour tous les pas de temps de la période d'étude, sans utilisation de l'information acquise au fil du temps, est un problème statique. Or, les commandes du $j^{\text{ème}}$ jour peuvent être logiquement influencées par la réalisation des aléas des jours antérieurs. Au jour j , nous connaissons ces réalisations θ_r^j . Nous avons donc à disposition de l'information non utilisée qui permettrait d'adapter les commandes. Pour remédier à cela, nous introduisons le concept de **recours**.

L'incorporation du recours dans la formulation du problème ne va pas sans difficulté. Si nous prenons une approche de type programmation stochastique, les décisions à chaque instant sont des recours, fonctions des aléas passés. Le fait d'avoir à optimiser sur un espace de fonctions très général conduit, pour les problèmes multi-étapes, à une explosion de la dimension. Pour rester dans un environnement numériquement traitable, nous nous proposons de restreindre l'espace des fonctions de recours à celui des fonctions linéaires des aléas observés. Nous parlons alors de **règle de décision linéaire**.

Cette restriction nous permet de modéliser le problème de départ dans le cadre de l'optimisation robuste comme un problème de programmation linéaire, certes de grande taille, mais parfaitement traitable par des solveurs numériques actuellement disponibles. Le prix à payer pour cette simplification considérable est, bien entendu, une certaine perte d'optimalité de la solution. Mais en l'absence de techniques et d'outils permettant de calculer une solution optimale, cette restriction peut se justifier.

Dans le cadre du problème hydraulique, la règle de décision linéaire $\tilde{F}_g^t(\theta)$ choisie peut se formuler ainsi :

$$\tilde{F}_g^t(\theta) = \sum_j \sum_r x_r^j \theta_r^j. \quad (4.4)$$

Les nouvelles variables du problème sont désormais les x_r^j . Ainsi, plus nous insérons de l'information dans la règle, plus le nombre de variables du problème croît.

Remarque 4.2 *Nous noterons que les recours sont dynamiques, mais que les paramètres x_r^j sont fixés en première période et sont donc statiques.*

4.3.3 Optimisation, incertitude et validation

Le but des sections suivantes est de présenter une approche opérationnelle de la gestion d'une vallée hydraulique qui prenne en compte explicitement l'incertitude sur les apports d'eau. Pour

mesurer l'intérêt de cette approche, il faut pouvoir la comparer à d'autres politiques de gestion. Nous en retenons trois. Le premier IP est un idéal qui ne peut pas être mis en œuvre, mais qui donne une borne inférieure au coût de gestion. La deuxième DF est simple, mais comme nous le verrons, irréaliste. La troisième DRP semble assez proche de ce qui est réalisé en pratique. Ces politiques sont détaillées ci-dessous :

La Politique à Information Parfaite (IP) Dans cette approche, nous considérons que les scénarios d'apport d'eau sont connus avant le premier pas de temps. L'optimisation s'effectue en avenir certain sur chacun des scénarios individuels. Cette politique permet de satisfaire exactement toutes les contraintes. Elle ne peut pas être mise en œuvre puisque les apports d'eau ne sont connus que progressivement. Il s'agit d'une politique idéale fournissant une borne inférieure aux coûts de gestion.

Politique déterministe fixe (DF) Nous considérons une version déterministe du problème dans laquelle la chronique des apports d'eau est celle des apports moyens. La chronique résultante des décisions de turbinage est mise en œuvre sur des scénarios tirés au hasard.

Déterministe à Révisions Périodiques (DRP) Dans cette approche, un planning est établi sur base d'une projection déterministe des apports d'eau moyens. Ce modèle est utilisé par l'exploitant pour gérer la première journée, avec au besoin une adaptation en temps réel des turbinés pour tenir compte des fluctuations infra-journalières. Au terme de la première journée, un nouveau planning basé sur les niveaux actualisés des réservoirs est calculé pour la période restante. Il est utilisé par l'exploitant comme précédemment. Cette politique de gestion est dit à « révisions périodiques ». La révision périodique du planning introduit une capacité d'adaptation indispensable. Le planning sur les jours restants n'est pas utilisé directement par l'exploitant mais est utile pour garantir que le plan de gestion du premier jour s'inscrit dans une perspective cohérente d'anticipation des prix et des besoins.

Ces politiques serviront de référence pour évaluer les mérites des deux approches robustes que nous proposons :

Robuste à Révisions Périodiques (RRP) Nous procédons aux mêmes opérations que pour la politique DRP à part que l'optimisation n'est plus déterministe mais robuste sur la première journée. La planification pour les jours suivants se fait sur un mode déterministe. Le but est d'introduire dans les décisions du premier jour une certaine vision du futur, sachant que la partie du planning au delà du premier jour sera remise à plat lors de la révision au jour suivant. En résumé, l'optimisation robuste doit assurer une protection contre des violations de contraintes dans la première journée ; la révision, comme précédemment, permet une adaptation.

Robuste à Ajustements Linéaires (RAL) L'approche RAL est plus ambitieuse que l'approche RRP car elle s'efforce d'introduire dans le modèle lui-même la capacité d'adaptation par le biais de règles de décision linéaires sur les turbinages. Ces règles donnent le moyen de gérer le système hydraulique sur toute la période de planification sans réactualisation journalière du modèle. Dans la pratique, un exploitant souhaitera toujours une réactualisation. Il faudrait donc implémenter RAL dans un contexte à révision périodique. Nous ne l'avons pas fait pour des raisons de temps calcul dans les simulations. L'intérêt de cette politique est de voir si les règles de décision linéaires permettent une adaptation efficace et surmontent leur handicap par rapport aux politiques DRP et RRP.

4.4 Politique de gestion Robuste à Révisions Périodiques (RRP) du problème hydraulique

Dans cette section (et la suivante), nous appliquons les résultats du chapitre 3 pour formaliser la politique de gestion RRP. Nous rappelons que dans cette politique, il n'est pas fait appel à des règles de décision linéaires et que seules les contraintes du premier jour sont soumises à une exigence de robustesse. Le mode d'exploitation du modèle se fait par révision périodique. Cela implique que les décisions des jours deux et suivants calculées par le modèle ne sont pas implémentées.

Les aléas considérés sont les aléas instantanés en chaque réservoir ϵ_r^t . Nous choisissons un ensemble d'incertitude $\ell_1 \cap \ell_\infty$ comme dans l'exemple précédent. L'ensemble s'écrit :

$$\Theta^t = \{\epsilon_r^\tau \mid \sum_{\tau=1}^t |\epsilon_r^\tau| \leq \kappa\sqrt{t}, |\epsilon_r^\tau| \leq \kappa, \forall r \in 1, \dots, R, \forall \tau \in 1, \dots, t\}$$

ou

$$\begin{aligned} \Theta^t = \{ & \epsilon_r^\tau \mid \epsilon_r^\tau = (\epsilon_r^\tau)^+ - (\epsilon_r^\tau)^-, \sum_{\tau=1}^t |\epsilon_r^\tau| \leq \kappa\sqrt{t}, \\ & |\epsilon_r^\tau| \leq \kappa, (\epsilon_r^\tau)^+ \geq 0, (\epsilon_r^\tau)^- \geq 0, \forall r \in 1, \dots, R, \forall \tau \in 1, \dots, t\}. \end{aligned}$$

L'ensemble d'incertitude dépend de t et s'agrandit en fonction de l'information disponible.

Nous rappelons que l'ensemble $\ell_1 \cap \ell_\infty$ nous sert à approximer une boule ℓ_2 de dimension n . Pour calibrer le rayon de la boule κ et définir une probabilité de présence dans cet ensemble, nous calculons le rayon à l'aide d'une loi χ_n^2 à n degrés de liberté (justifié en tant que somme des carrés de n aléas centrés réduits indépendants).

Puisque les valeurs réalisées des aléas θ_r^1 de première journée ne sont révélées qu'à la fin de cette journée, les commandes ne sont pas des recours. Elles sont déterministes et doivent être fixées dès l'instant initial. Les contraintes à rendre robustes sont celles sur les volumes $\underline{V}_r \leq \tilde{V}_r^t \leq \bar{V}_r^t$.

Nous détaillons uniquement l'équivalent robuste de la première contrainte. L'autre contrainte se déduit facilement. Nous proposons de rendre ces contraintes robustes vis-à-vis des aléas d'apports instantanés ϵ_r^τ . La première des deux contraintes s'écrit :

$$V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sigma_r \epsilon_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} \geq \underline{V}_r^t.$$

L'équivalent robuste s'obtient en résolvant le problème

$$\min_{\epsilon} \left\{ \sum_{\tau \leq t} \sigma_r \epsilon_r^\tau \mid \sum_{\tau \leq t} |\epsilon_r^\tau| \leq \kappa\sqrt{t}, |\epsilon_r^\tau| \leq \kappa, \forall r \in 1, \dots, R, \forall \tau \in 1, \dots, t \right\}.$$

ou

$$\min_{\epsilon} \left\{ \sum_{\tau \leq t} \sigma_r [(\epsilon_r^\tau)^+ - (\epsilon_r^\tau)^-] \right\}$$

4.5. POLITIQUE DE GESTION ROBUSTE À AJUSTEMENTS LINÉAIRES (RAL) DU PROBLÈME HYDRAULIQUE

sous les contraintes :

$$\begin{aligned} \sum_{\tau=1}^t [(\epsilon_r^\tau)^+ + (\epsilon_r^\tau)^-] &\leq \kappa\sqrt{t}, \quad \forall r \in 1, \dots, R \\ 0 \leq (\epsilon_r^\tau)^+ &\leq \kappa, \quad \forall r \in 1, \dots, R, \quad \forall \tau \in 1, \dots, t \\ 0 \leq (\epsilon_r^\tau)^- &\leq \kappa, \quad \forall r \in 1, \dots, R, \quad \forall \tau \in 1, \dots, t \end{aligned}$$

En passant au dual, l'équivalent robuste de la contrainte de volume minimal s'écrit :

$$\begin{aligned} V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} - v_{r,t} \kappa \sqrt{t} - \kappa \sum_{\tau \leq t} (u_{r,\tau}^t)_p - \kappa \sum_{\tau \leq t} (u_{r,\tau}^t)_m &\geq \underline{V}_r^t \\ (u_{r,\tau}^t)_p + v_{r,t} &\geq -\sigma_r, \quad \forall t \in 1, \dots, T, \quad \forall \tau \in 1, \dots, t, \quad \forall r \in 1, \dots, R \\ (u_{r,\tau}^t)_m + v_{r,t} &\geq \sigma_r, \quad \forall t \in 1, \dots, T, \quad \forall \tau \in 1, \dots, t, \quad \forall r \in 1, \dots, R \\ (u_{r,\tau}^t)_p &\geq 0, \quad (u_{r,\tau}^t)_m \geq 0, \quad v_{r,t} \geq 0. \end{aligned}$$

De même, l'équivalent robuste de la contrainte supérieure de volume s'écrit :

$$\begin{aligned} V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} + v_{r,t} \kappa \sqrt{t} + \kappa \sum_{\tau \leq t} (u_{r,\tau}^t)_p + \kappa \sum_{\tau \leq t} (u_{r,\tau}^t)_m &\leq \bar{V}_r^t \\ (u_{r,\tau}^t)_p + v_{r,t} &\geq \sigma_r, \quad \forall t \in 1, \dots, T, \quad \forall \tau \in 1, \dots, t, \quad \forall r \in 1, \dots, R \\ (u_{r,\tau}^t)_m + v_{r,t} &\geq -\sigma_r, \quad \forall t \in 1, \dots, T, \quad \forall \tau \in 1, \dots, t, \quad \forall r \in 1, \dots, R \\ (u_{r,\tau}^t)_p &\geq 0, \quad (u_{r,\tau}^t)_m \geq 0, \quad v_{r,t} \geq 0. \end{aligned}$$

Pour les pas de temps suivant T , les contraintes de volume considérées sont les contraintes déterministes classiques (et donc uniquement basées sur les prévisions d'apports) :

$$\underline{V}_r^t \leq V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} \leq \bar{V}_r^t, \quad \forall r \in 1, \dots, R, \quad \forall t \in T+1, \dots, H.$$

L'accroissement de la taille du modèle robuste par rapport au modèle déterministe est contrôlé : nous ne rajoutons que $2RT(T+1) + 2RT$ variables et $2RT(T+1)$ contraintes.

4.5 Politique de gestion Robuste à Ajustements Linéaires (RAL) du problème hydraulique

Contrairement à la section 4.4, nous nous plaçons dans la situation où la politique de turbinage est déterminée à un instant initial sur tout l'ensemble de la période d'étude. Ne bénéficiant

4.5. POLITIQUE DE GESTION ROBUSTE À AJUSTEMENTS LINÉAIRES (RAL) DU PROBLÈME HYDRAULIQUE

donc plus des ré-optimisations successives de la politique RRP, le modèle doit donc comporter un mécanisme d'auto-adaptation qui prend en compte les réalisations des aléas passés. Pour cela, nous introduisons les recours sous forme de règles de décision linéaires et nous formulons l'équivalent robuste des contraintes pour tous les pas de temps de planification et non plus uniquement sur ceux de la première journée. Le modèle obtenu est plus complexe et plus grand que le modèle RRP.

Comme dans la section précédente, nous choisissons un ensemble d'incertitude $\ell_1 \cap \ell_\infty$ pour les bruits blancs θ :

$$\Theta^t = \{\theta_r^j \mid \sum_{j=1}^{J(t)} \sum_{r=1}^R |\theta_r^j| \leq \kappa \sqrt{J(t)R}, |\theta_r^j| \leq \kappa, \forall r \in 1, \dots, R, \forall j \in 1, \dots, J(t)\}$$

ou

$$\begin{aligned} \Theta^t = \{ \theta_r^j \mid \theta_r^j &= (\theta_r^j)^+ - (\theta_r^j)^-, \sum_{j=1}^{J(t)} \sum_{r=1}^R |\theta_r^j| \leq \kappa \sqrt{J(t)R}, \\ &|\theta_r^j| \leq \kappa, (\theta_r^j)^+ \geq 0, (\theta_r^j)^- \geq 0, \forall r \in 1, \dots, R, \forall j \in 1, \dots, J(t) \}. \end{aligned}$$

Nous ne considérons plus les aléas sur les pas de temps individuels. Dans cette formulation, l'évolution des apports d'eau durant la journée est traitée comme si elle était déterministe, ce qui permet de simplifier le modèle. Pour rendre RAL comparable à RRP, nous pouvons rendre robustes les contraintes du premier jour vis-à-vis des aléas des pas de temps individuels de ce jour. Nous obtenons ainsi deux variantes de RAL : la première RALD (D pour déterministe) ignore les aléas par pas de temps de la première journée ; la deuxième RALR (R pour robuste) en tient compte. Une troisième version pourrait prendre en compte tous les aléas individuels sur les 84 pas de temps. Ce serait plus réaliste, mais conduirait à un modèle de trop grande taille.

La règle choisie fait dépendre la décision en un pas de temps des aléas agrégés des k derniers jours pour tous les réservoirs. Elle s'écrit donc :

$$\begin{aligned} \tilde{F}_{gt} &= F_{gt} + \sum_{l=1}^k x_{gt}^l \sum_{r=1}^R \theta_r^{J(t)-l+1} \\ &= F_{gt} + \sum_{l=1}^k x_{gt}^l \theta^{J(t)-l+1}. \end{aligned} \tag{4.5}$$

Différentes règles ont été testées (dont une plus complexe avec un cumul de tous les apports antérieurs) mais nous ne montrerons dans cette étude que l'application de la règle (4.5) qui a fourni les meilleurs résultats numériques.

Nous pouvons maintenant écrire formellement l'équivalent robuste des contraintes du problème.

4.5. POLITIQUE DE GESTION ROBUSTE À AJUSTEMENTS LINÉAIRES (RAL) DU PROBLÈME HYDRAULIQUE

4.5.1 Contraintes de turbinage

Les contraintes sur les quantités turbinées sont

$$\underline{F}_{gt} \leq \tilde{F}_{gt} \leq \bar{F}_{gt}.$$

Nous détaillons uniquement l'obtention de l'équivalent robuste de la contrainte de turbinage minimal. L'autre contrainte se déduit facilement. La contrainte robuste s'écrit comme

$$F_{gt} + \sum_{l=1}^k x_{gt}^l \theta^{J(t)-l+1} \geq \underline{F}_{gt}, \quad \text{pour tout } \theta \in \Theta^{t,k}$$

où $\Theta^{t,k}$ est l'ensemble d'incertitude qui dépend du type de règle retenu.

La forme de l'ensemble d'incertitude dépend de la valeur relative de $J(t)$ par rapport à k . Nous avons comme expression :

$$\begin{array}{ll} \sum_{j=J(t)-k+1}^{J(t)} |\theta^j| \leq \kappa \sqrt{kR} & \text{si } J(t) \geq k \\ \vdots & \vdots \\ \sum_{j=1}^{J(t)} |\theta^j| \leq \kappa \sqrt{J(t)R} & \text{si } 1 \leq J(t) < k. \end{array}$$

Nous pouvons généraliser cette écriture par :

$$\sum_{j=J(t)-\min(J(t),k)+1}^{J(t)} |\theta_j| \leq \kappa \sqrt{\min(J(t),k) \cdot R}, \quad \forall j \in 1, \dots, J(t).$$

Posons :

$$k' = \min(J(t), k).$$

L'équivalent robuste s'obtient donc en résolvant le problème

$$\min \left\{ \sum_{l=1}^{k'} x_{gt}^l \theta^{J(t)-l+1} \mid \sum_{j=J(t)-k'+1}^{J(t)} |\theta^j| \leq \kappa \sqrt{k'R}, |\theta^j| \leq \kappa, \forall j \in \{J(t) - k' + 1, \dots, J(t)\} \right\}.$$

Nous supprimons les valeurs absolues à l'aide des variables positives $(\theta^j)^+$ et $(\theta^j)^-$. Le problème primal à résoudre est donc :

$$\min_{\theta} \left\{ \sum_{l=1}^{k'} x_{gt}^l \sum_{r=1}^R [(\theta_r^{J(t)-l+1})^+ - (\theta_r^{J(t)-l+1})^-] \right\}$$

sous les contraintes :

$$\begin{aligned} & \sum_{j=J(t)-k'+1}^{J(t)} \sum_{r=1}^R [(\theta_r^j)^+ + (\theta_r^j)^-] \leq \kappa \sqrt{k'R} \\ & 0 \leq (\theta_r^j)^+ \leq \kappa, \quad \forall j \in \{J(t) - k' + 1, \dots, J(t)\}, \quad \forall r \in 1, \dots, R \\ & 0 \leq (\theta_r^j)^- \leq \kappa, \quad \forall j \in \{J(t) - k' + 1, \dots, J(t)\}, \quad \forall r \in 1, \dots, R \end{aligned}$$

4.5. POLITIQUE DE GESTION ROBUSTE À AJUSTEMENTS LINÉAIRES (RAL) DU PROBLÈME HYDRAULIQUE

En introduisant les multiplicateurs $v_{g,t}$, $(u_{g,t}^{r,j})_p$ et $(u_{g,t}^{r,j})_m$ associés aux précédentes contraintes, le problème dual associé au primal précédent est donc :

$$\max\{-v_{g,t}\kappa\sqrt{k'R} - \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_p - \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_m\}$$

sous

$$\begin{aligned} (u_{g,t}^{r,j})_p + v_{g,t} &\geq -x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_m + v_{g,t} &\geq x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_p &\geq 0, (u_{g,t}^{r,j})_m \geq 0, v_{g,t} \geq 0 \end{aligned}$$

L'équivalent robuste de la contrainte de turbinage inférieur s'écrit donc :

$$F_{gt} - \kappa v_{g,t}\sqrt{k'R} - \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_p - \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_m \geq \underline{F}_{gt}$$

$$\begin{aligned} (u_{g,t}^{r,j})_p + v_{g,t} &\geq -x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_m + v_{g,t} &\geq x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_p &\geq 0, (u_{g,t}^{r,j})_m \geq 0, v_{g,t} \geq 0 \end{aligned}$$

De même, l'équivalent robuste de la contrainte supérieure de turbinage s'écrit :

$$F_{gt} + v_{g,t}\kappa\sqrt{k'R} + \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_p + \kappa \sum_{r=1}^R \sum_{j=J(t)-k'+1}^{J(t)} (u_{g,t}^{r,j})_m \leq \overline{F}_{gt}$$

$$\begin{aligned} (u_{g,t}^{r,j})_p + v_{g,t} &\geq x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_m + v_{g,t} &\geq -x_{gt}^l, \quad \forall l \in 1, \dots, k', \quad j = J(t) - l + 1 \\ (u_{g,t}^{r,j})_p &\geq 0, (u_{g,t}^{r,j})_m \geq 0, v_{g,t} \geq 0 \end{aligned}$$

4.5.2 Contraintes de volume

Les contraintes sur les volumes sont :

$$\underline{V}_r^t \leq \tilde{V}_r^t \leq \overline{V}_r^t.$$

La contrainte de volume minimal s'écrit :

$$V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{j \leq J(t)} \sigma_r \sqrt{T} \theta_r^j + \sum_{t=J(t).T+1}^t \sigma_r \epsilon_r^t + \sum_{\tau \leq t} \sum_g \delta_{gr} \tilde{F}_{g\tau} \geq \underline{V}_r^t.$$

4.5. POLITIQUE DE GESTION ROBUSTE À AJUSTEMENTS LINÉAIRES (RAL) DU PROBLÈME HYDRAULIQUE

Nous détaillons l'équivalent robuste de la première contrainte seulement. L'autre contrainte se déduit facilement. Pour écrire la contrainte robuste, explicitons l'expression du volume :

$$\begin{aligned}\tilde{V}_r^t = V_r^0 &+ \sum_{\tau \leq t} a_r^\tau + \sum_{j \leq J(t)} \sigma_r \sqrt{T} \theta_r^j + \sum_{\tau = J(t).T+1}^t \sigma_r \epsilon_r^\tau \\ &+ \sum_{\tau \leq t} \sum_g \delta_{gr} (F_{g\tau} + \sum_{l=1}^{k'} x_{g\tau}^l \theta^{J(\tau)-l+1}).\end{aligned}$$

Nous proposons de rendre ces contraintes robustes vis-à-vis des aléas θ_r^j , mais nous négligeons les aléas d'apports journaliers ϵ_r^j . Notons que cette opération se fera par référence à des variables θ_r^j qui n'interviennent pas directement dans la règle de décision. Cela rend l'équivalent robuste des contraintes de volume plus complexe que ce que nous avons obtenu pour les contraintes de turbinage.

L'équation du volume peut se mettre sous la forme condensée

$$\tilde{V}_r^t = V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} + \sum_{j \leq J(t)} \sum_{r'=1}^R \alpha_{r'r}^{tj} \theta_{r'}^j,$$

avec

$$\begin{aligned}\alpha_{r'r}^{tj} &= I_{r'r} \sigma_r \sqrt{T} + \sum_g \delta_{gr} \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,t)} x_{g\tau}^l \right), \\ s(l,j,t) &= \min\{(j+l)T, t\}, \\ I_{r'r} &= \begin{cases} 1 & \text{si } r' = r, \\ 0 & \text{autrement.} \end{cases}\end{aligned}$$

L'équivalent robuste s'obtient en résolvant le problème

$$\min_{\theta} \left\{ \sum_{r'=1}^R \sum_{j \leq J(t)} \alpha_{r'r}^{tj} \theta_{r'}^j \mid \sum_{j=1}^{J(t)} \sum_{r'=1}^R |\theta_{r'}^j| \leq \kappa \sqrt{J(t)R}, |\theta_{r'}^j| \leq \kappa, \forall j \in 1, \dots, J(t), \forall r' \in 1, \dots, R \right\}.$$

ou

$$\min_{\theta^+, \theta^-} \left\{ \sum_{r'=1}^R \sum_{j \leq J(t)} \alpha_{r'r}^{tj} [(\theta_{r'}^j)^+ - (\theta_{r'}^j)^-] \right\}$$

sous

$$\begin{aligned}\sum_{j=1}^{J(t)} \sum_{r'=1}^R [(\theta_{r'}^j)^+ + (\theta_{r'}^j)^-] &\leq \kappa \sqrt{J(t)R} \\ 0 \leq (\theta_{r'}^j)^+ &\leq \kappa, \forall j \in 1, \dots, J(t), \forall r' \in 1, \dots, R \\ 0 \leq (\theta_{r'}^j)^- &\leq \kappa, \forall j \in 1, \dots, J(t), \forall r' \in 1, \dots, R\end{aligned}$$

En passant au dual, l'équivalent robuste de la contrainte de volume minimal s'écrit :

$$V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} - v_{r,t} \kappa \sqrt{J(t)R} - \kappa \sum_{r'=1}^R \sum_{j \leq J(t)} (u_{r,t}^{r',j})_p - \kappa \sum_{r'=1}^R \sum_{j \leq J(t)} (u_{r,t}^{r',j})_m \geq \underline{V}_r^t$$

$$\begin{aligned} (u_{r,t}^{r',j})_p + v_{r,t} &\geq -\alpha_{r'r}^{tj}, \quad \forall j \in 1, \dots, J(t), \quad \forall r' \in 1, \dots, R \\ (u_{r,t}^{r',j})_m + v_{r,t} &\geq \alpha_{r'r}^{tj}, \quad \forall j \in 1, \dots, J(t), \quad \forall r' \in 1, \dots, R \\ (u_{r,t}^{r',j})_p &\geq 0, \quad (u_{r,t}^{r',j})_m \geq 0, \quad v_{r,t} \geq 0. \end{aligned}$$

De même, l'équivalent robuste de la contrainte supérieure de volume s'écrit :

$$V_r^0 + \sum_{\tau \leq t} a_r^\tau + \sum_{\tau \leq t} \sum_g \delta_{gr} F_{g\tau} + v_{r,t} \kappa \sqrt{J(t)R} + \kappa \sum_{r'=1}^R \sum_{j \leq J(t)} (u_{r,t}^{r',j})_p + \kappa \sum_{r'=1}^R \sum_{j \leq J(t)} (u_{r,t}^{r',j})_m \leq \overline{V}_r^t$$

$$\begin{aligned} (u_{r,t}^{r',j})_p + v_{r,t} &\geq \alpha_{r'r}^{tj}, \quad \forall j \in 1, \dots, J(t), \quad \forall r' \in 1, \dots, R \\ (u_{r,t}^{r',j})_m + v_{r,t} &\geq -\alpha_{r'r}^{tj}, \quad \forall j \in 1, \dots, J(t), \quad \forall r' \in 1, \dots, R \\ (u_{r,t}^{r',j})_p &\geq 0, \quad (u_{r,t}^{r',j})_m \geq 0, \quad v_{r,t} \geq 0 \end{aligned}$$

4.5.3 Fonction-objectif

Contrairement aux contraintes physiques précédentes où aucune violation n'est autorisée, nous pouvons nous poser la question de rendre robuste l'objectif. En effet, si nous souhaitons garantir un coût minimal, nous pouvons transformer l'objectif en une contrainte d'inégalité que nous pouvons rendre robuste. Ici, l'objectif étant d'avoir des commandes robustes au coût le moins élevé possible, il semble opportun de ne pas robustifier la fonction-objectif (ce qui est confirmé avec les résultats numériques). Cependant, la méthode d'obtention de l'équivalent robuste dans cette étude est tout de même fournie en Annexe A.

4.6 Tailles des modèles

Une expression analytique de la taille des divers modèles proposés peut être établie.

Nous proposons ci-dessous d'appliquer ces formules au cas suivant : le pas de temps est bi-horaire et la règle de décision ne porte que sur le dernier aléa. Pour la commodité, nous donnons ces expressions sur trois tableaux séparés Tab. 4.1, Tab. 4.2 et Tab. 4.4, un pour chaque modèle.

4.6. TAILLES DES MODÈLES

Variables	
Commandes (turbinage)	$G \times H$
Contraintes	
Commandes (bornes sur les turbinages)	$2G \times H$
Volumes	$2R \times H$
Objectif	1

TAB. 4.1 – Taille du modèle déterministe

Variables	
Commandes (turbinage)	$G \times H$
Variables auxiliaires $u_{r\tau}^{t+}$ et $u_{r\tau}^{t-}$	$2R \times T \times (T + 1)$
Variables auxiliaires v_{rt}	$2R \times T$
Contraintes	
Commandes (bornes sur les turbinages)	$2G \times T$
Limites sur les volumes	$2R \times H$
Contraintes auxiliaires sur les volumes	$2R \times T \times (T + 1)$
Objectif	1

TAB. 4.2 – Taille du modèle Robuste à Révisions Périodiques RRP

Le premier exemple considéré est le jeu d’essais utilisé dans la section “Résultats” de cette note. Nous considérons 3 réservoirs, 6 groupes et 84 pas de temps. Cet exemple est certes de petite taille mais représentatif de plusieurs vallées du parc EDF. La taille des modèles est donnée sur le tableau (Tab. 4.3).

Type	RAL	RRP	Déterministe
Nombre de variables	18 812	1 513	505
Nombre de contraintes	18 092	2 450	1 514

TAB. 4.3 – Modèles Déterministe, RAL et RRP avec 3 réservoirs, 6 groupes et 84 pas de temps

Considérons maintenant une vallée bien plus importante de 15 réservoirs, 60 groupes sur 144 pas de temps. Une telle vallée est représentative des plus grandes vallées EDF. La taille des modèles est donnée sur le tableau (Tab. 4.5).

4.6. TAILLES DES MODÈLES

Variables	
Commandes de turbinage	
Variables $Turb_g^t$	$G \times H$
Variables (coeff. de la règle de décision) x_{gt}^l	$G \times T \times K \times (J(H) - (K/2) + 0.5)$
Variables auxiliaires u_{gt}^{rj+} et u_{gt}^{rj-}	$2R \times G \times T \times K \times (2J(H) - K + 1)$
Variables v_{gt}	$2G \times J(H) \times T$
Volumes	
Variables auxiliaires u_{gt}^{rj+} et u_{gt}^{rj-}	$2R^2 \times T \times J(H) \times (J(H) + 1)$
Variables auxiliaires v_{gt}	$2R \times J(H) \times T$
Variables $\alpha_{r'r}^{tj}$	$(R^2 \times T)/2 \times J(H) \times (J(H) + 1)$
Objectif	
Variables auxiliaires u_{rj}^+ et u_{rj}^-	$2R \times J(H)$
Variable auxiliaire v	1
Variables auxiliaires $\beta_{r'}^j$	$R \times J(H)$
Contraintes	
Turbinages	
Commandes (bornes sur les turbinages)	$2G \times T$
Contraintes auxiliaires sur les commandes	$2R \times G \times T \times K \times (2J(H) - K + 1)$
Volumes	
Limites sur les volumes	$2R \times H$
Contraintes auxiliaires sur les volumes	$2R^2 \times T \times J(H) \times (J(H) + 1)$
Contraintes auxiliaires de définition des $\alpha_{r'r}^{tj}$	$(R^2 \times T)/2 \times J(H) \times (J(H) + 1)$
Objectif	
Objectif	1
Contraintes auxiliaires pour l'objectif	$2R \times J(H)$
Contraintes de définition des $\beta_{r'}^j$	$R \times J(H)$

TAB. 4.4 – Taille du modèle Robuste à Ajustements Linéaires RAL

Type	RAL	RRP	Déterministe
Nombre de variables	566 732	13 681	8 641
Nombre de contraintes	564 572	26 282	21 602

TAB. 4.5 – Modèles Déterministe, RAL et RRP avec 15 réservoirs, 60 groupes sur 144 pas de temps

La différence entre les tailles des modèles Déterministe et RRP n'est pas vraiment significative. D'autre part, l'équivalent robuste RAL correspondant est certes grand (567 000 variables et 565 000 contraintes) mais peut être résolu par des codes de programmation linéaire performants. Cet accroissement de la taille des modèles n'est donc pas une entrave à l'application d'approches robustes.

4.7 Etude expérimentale

4.7.1 Implémentation

Le problème hydraulique robuste, bien que linéaire, a une écriture complexe qui rend très difficile la construction de la matrice des contraintes. Surtout si le nombre de contraintes est très important. Or, cette matrice est primordiale comme donnée d'entrée de solveurs d'optimisation comme CPLEX, Express,... Pour remédier à ce problème, nous avons utilisé un **langage de modélisation**. Avec un tel outil, l'utilisateur n'a plus qu'à créer un fichier de données, définir son problème avec une écriture très proche de l'écriture mathématique et le langage de modélisation se charge de la transformation du problème et du lien au solveur. Ainsi, avec quelques bases de programmation, nous pouvons facilement modéliser le problème et le résoudre. Un tel outil est vraiment très pratique et permet de faire des essais de modèles rapidement et très simplement.

Tout au cours de ce travail, nous avons utilisé le "GNU Linear Programming Kit" (GLPK), composé du langage de programmation GNU MathProg et du solveur glpsol. GNU MathProg est un sous-ensemble du langage de modélisation AMPL. AMPL est un langage de modélisation très puissant que nous pouvons coupler avec divers solveurs comme CPLEX, EXPRESS, MOSEK,... glpsol est un solveur linéaire MIP limitant le kit aux problèmes linéaires en variables réelles et/ou entières. Notons qu'AMPL, couplé au solveur adéquat, peut être facilement utilisé à des problèmes plus complexes (optimisation non linéaire, optimisation conique, ...).

Nous avons créé une maquette en langage MATLAB permettant :

1. la définition du jeu d'essais (Topologie de la vallée, nombre de pas de temps,...) ;
2. la construction du fichier de données `.dat` ;
3. l'optimisation du problème avec appel de GLPK dans Matlab ;
4. la génération de scénarios d'apports ;
5. l'application de la solution obtenue sur l'ensemble des scénarios ;

6. le calcul d'indicateurs : coût moyen, le nombre et la fréquence des violations de contraintes,...

Nous pouvons aussi utiliser cette maquette en horizon glissant.

4.7.2 Jeu d'essais

La vallée considérée est inspirée du début de la vallée de l'Ain. Elle est constituée de 3 réservoirs et 6 groupes sur un horizon de 84 pas de temps de 2 heures (soit 7 jours). Pour passer d'une unité de m^3/s à une unité de m^3/pdt (pdt : pas de temps), la taille des réservoirs a été divisée par la durée du pas de temps soit 7 200s. Nous considérons trois types de réservoir : un grand réservoir de tête, valorisé avec une valeur de l'eau, un réservoir intermédiaire jouant le rôle d'une écluse (la quantité d'eau reçue peut être évacuée sur la période d'étude) et un réservoir final de taille inférieure au réservoir de tête. La situation initiale est une situation peu contrainte puisque les volumes initiaux de chaque réservoir sont relativement éloignés des bornes V_{min} et V_{max} .

Enfin, la courbe de coûts marginaux choisie est la courbe réelle d'une semaine de Janvier 2006 (mercredi à mardi). Les coûts marginaux étant relativement plus grands que les valeurs de l'eau, nous nous attendons à ce que les groupes turbinent à plein régime très fréquemment.

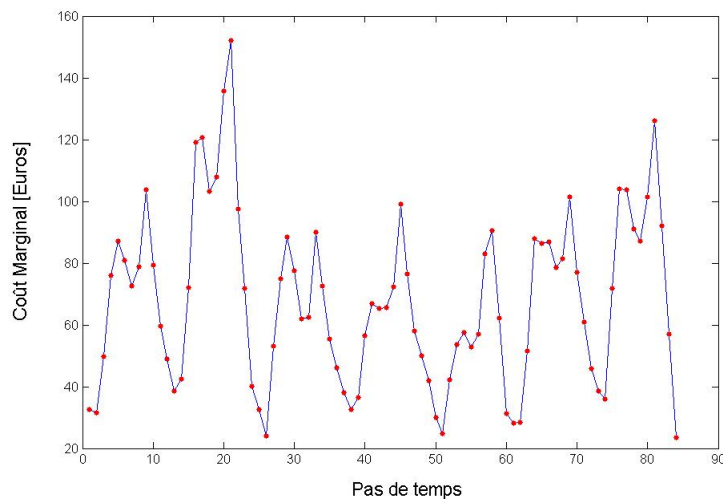


FIG. 4.2 – Chronique de Coûts Marginaux

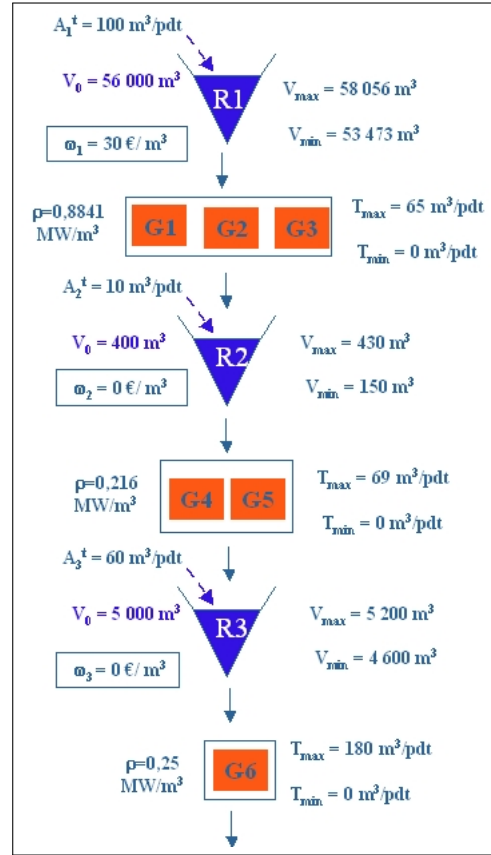


FIG. 4.3 – Vallée considérée

4.7.3 Simulation

Principe de la simulation

Une fois la phase d'optimisation achevée (qu'elle soit déterministe ou robuste), les commandes de turbinage obtenues doivent être testées pour valider l'efficacité de cette solution. Nous générerons dans ce but 10 000 scénarios d'apports hydrauliques selon une loi normale bi-tronquée sur lesquels nous appliquons les commandes optimisées de turbinage.

Remarque 4.3 Ces mêmes 10 000 scénarios seront utilisés dans chacune des études de la section suivante (Section 4.8).

L'application directe des commandes peut conduire à des violations de bornes sur les volumes minimal et maximal. A défaut de pouvoir simuler le comportement d'un exploitant réagissant en temps réel, nous pouvons essayer d'en donner un modèle plausible. Notre proposition consiste à définir les corrections de turbinés suivantes :

- Si les turbinés proposés conduisent à un dépassement de V_{\max} , nous ajustons ces turbinés de sorte que les violations disparaissent. Si les violations perdurent, l'eau en surplus est considérée comme du *déversement* et est déplacée vers le réservoir en aval.

- Quand le volume descend en dessous de V_{min} , les turbines sont stoppées jusqu'à ce que les apports au réservoir considéré (apports naturels, turbinés amont, déversés) comblent ces violations. Cette situation ne peut être autorisée que parce que $V_{min} > 0$ est déjà un niveau de protection.

Dans les simulations, nous constatons des déversés importants pour certaines politiques de gestion, très faibles pour d'autres, mais jamais absents. La question se pose de savoir si la contrainte sur les déversés est impérative ou si des violations sont permises exceptionnellement. Il semblerait que la deuxième possibilité prévaut dans la réalité. Il faudrait donc pouvoir associer à ces violations des pénalités qui permettent une comparaison effective entre les politiques de gestion. Dans la pratique, il est délicat de fournir une valeur à ces pénalités, si bien que nous nous contenterons de donner une information numérique sur l'intensité et la fréquence de ces violations.

Simulation du modèle DRP

Le schéma du processus de simulation du Déterministe à Révisions Périodiques peut être résumé comme suit :

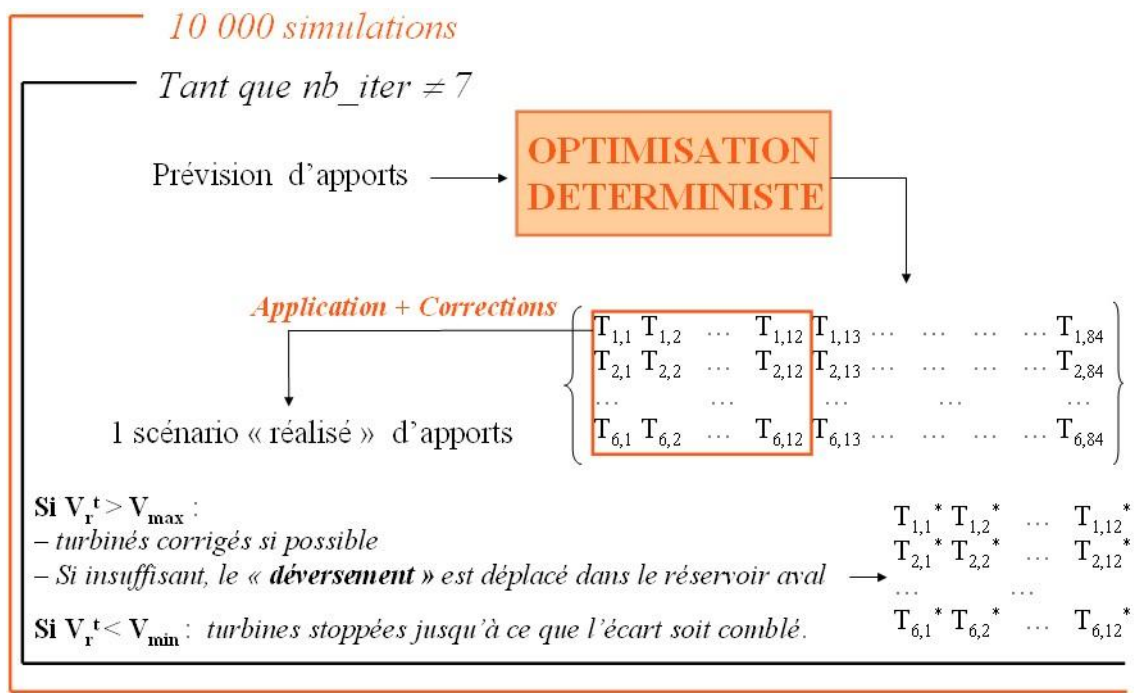


FIG. 4.4 – Processus de simulation du modèle DRP

Une révision est effectuée tous les jours, ce qui conduit à ne conserver que les commandes du premier jour à chaque optimisation. L'objectif de la révision est de donner à l'optimisation une condition initiale, les niveaux de stock, basée sur la réalisation des apports durant la journée. Pour cela, nous effectuons donc 7 optimisations sur chacun des 10 000 scénarios et les commandes

obtenues sont corrigées par l'heuristique de la sous-section 4.7.3 précédente. Finalement, nous calculons des indicateurs statistiques comme les coûts moyens, le cumul de déversés,...

Simulation des modèles robustes

Nous étudions deux types de modèle robuste. Le premier, Robuste à Révisions Périodiques (RRP), est similaire au DRP. Le schéma 4.5 le montre bien. En fait, la seule différence est que, dans le modèle d'optimisation, les commandes de turbinage du premier jour sont rendues robustes.

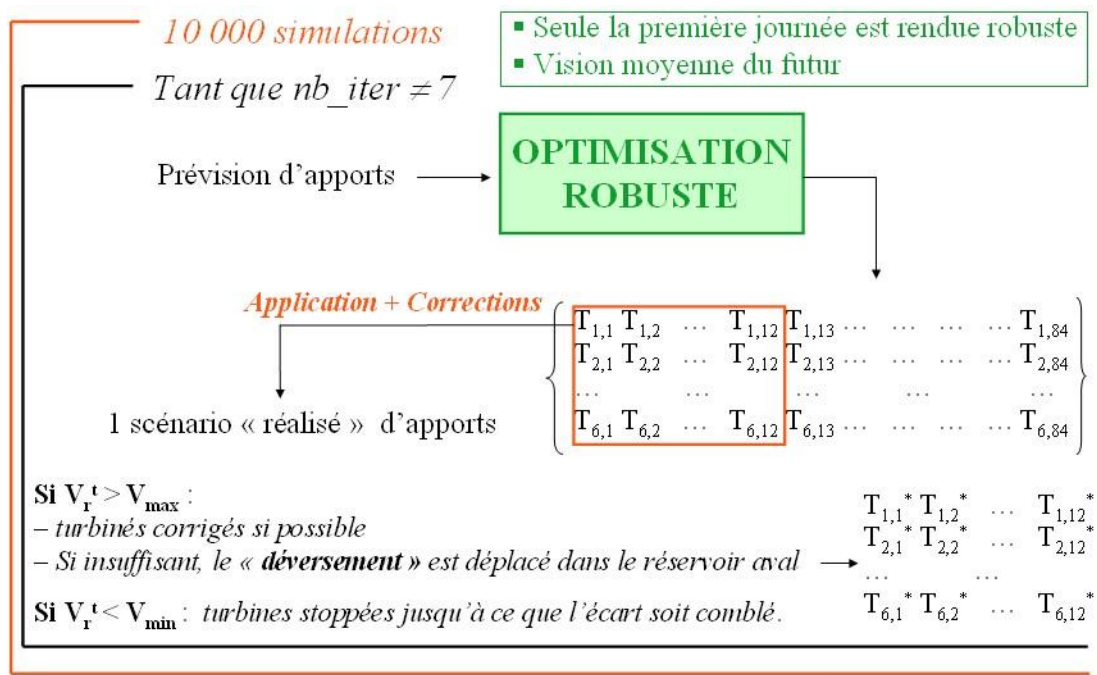


FIG. 4.5 – Processus de simulation du modèle RRP

Le deuxième type de modèle est le modèle Robuste à Ajustements Linéaires (RAL). Nous supposons donc prendre des décisions à l'instant initial pour tout l'horizon, sans révision (plus que 10 000 optimisations). En revanche, les commandes sont désormais des règles de décision. La solution obtenue par optimisation robuste s'écrit donc sous forme de coefficients F_{gt} et x_{gt}^l dans les règles de décision linéaires. L'interprétation de ces coefficients n'est pas immédiate. L'important toutefois est que ces coefficients définissent une stratégie de gestion opérationnelle qui peut être facilement mise en œuvre. Nous avons ainsi la possibilité d'en tester les mérites en procédant à des simulations. Dans l'exemple que nous traitons ici, nous procédons de la façon suivante. Pour chaque scénario tiré, nous reconstituons l'ensemble des commandes de turbinage ; nous les répartissons ensuite sur les groupes par rendement décroissant suivant l'heuristique décrite en 4.7.3. Nous appliquons ces commandes ajustées et nous mesurons l'évolution des niveaux d'eau obtenus (avec les adaptations explicitées auparavant) dans chacun des réservoirs. Nous calculons enfin les coûts afférents. Des indicateurs statistiques peuvent être calculés sur l'ensemble

des scénarios comme les coûts moyens et les écarts-types, les violations moyennes par scénario, les déversés moyens, etc.

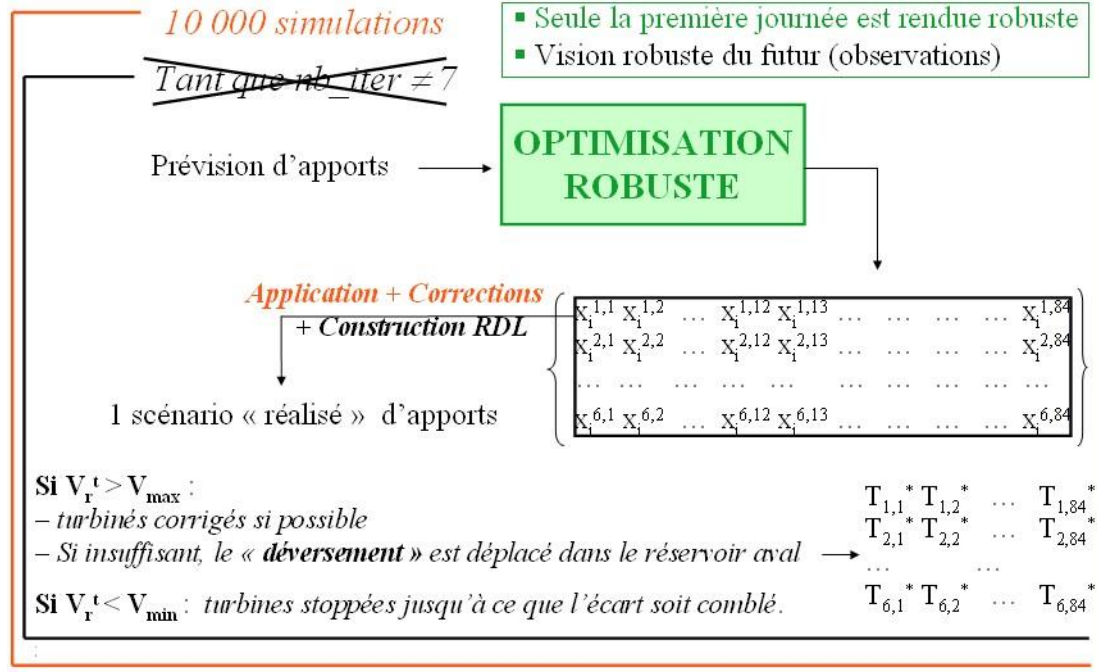


FIG. 4.6 – Processus de simulation du modèle RAL

4.8 Expérimentations numériques

Dans cette section, nous comparons numériquement les différentes approches définies en 4.3.3. Toutes ces approches (hormis IP) sont corrigées heuristiquement par 4.7.3. Dans tous les modèles robustes suivants, nous n'introduisons aucune robustesse dans l'objectif (voir annexe A pour la justification).

4.8.1 Comparaison entre DF et DRP

Nous tirons 10 000 scénarios d'apports à l'aide d'une loi normale bi-tronquée et nous comparons les coûts moyens, les violations moyennes sur les contraintes de volume et les déversés. Il est intéressant de comparer les politiques DF et DRP avec celui d'une gestion en information parfaite (IP). Le tableau 4.6 montre que, grâce à la révision journalière, DRP améliore DF sur les deux indicateurs "coût" et "violations". De plus DRP améliore le coût de IP. Ce résultat paradoxal s'explique par le fait que les violations de contrainte réalisées avec DF et DRP autorisent une amélioration des coûts de gestion ! Les informations de coût de DF et DRP ne sont donc pas pertinentes.

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

Politique de gestion	IP	DF	DRP
Coût moyen [euros]	-1 092 493	-1 092 579	-1 092 679
Déversé moyen [$\text{m}^3/\text{scén.}$]	0	76,2	43,5
Viol. moy. sur V_{min} [$\text{m}^3/\text{scén.}$]	0	55,5	7,4

TAB. 4.6 – Comparaison de DF et DRP avec une gestion en information parfaite (IP)

L'information du tableau 4.6 doit rester présente à l'esprit lorsque nous évaluons les performances des politiques de gestion que nous analyserons plus loin. Le respect des contraintes augmente à l'évidence les coûts de gestion.

Synthèse 4.1 (DF vs. DRP)

1. *DF entraîne des violations trop importantes pour pouvoir être mise en œuvre avec les ajustements simples des turbinés. L'espérance du coût qui lui est attachée n'est pas pertinente.*
2. *DRP permet une réduction des violations de contrainte, mais ces violations restent encore trop importantes pour rendre crédible la performance économique.*
3. *Le coût de IP peut seul être pris comme référence par la suite.*

4.8.2 Comparaison entre DRP et RRP

L'atout principal de l'approche RRP est sa protection contre les violations de contraintes de volume. Le tableau Tab.4.7 compare les déversés (violations en V_{max}) et les violations en V_{min} entre DRP et RRP (avec une couverture de 90%).

Techniques	Déversé moy. [$\text{m}^3/\text{scénario}$]	Pire scénario de déversés cumulés [m^3]	Déversé instant. maximal [m^3]	Nb moy. de pdt avec déversés [pas de tps/scén.]	Nbre de scén. avec déversés [scénarios]
Politique DRP	43,5	163,5	37,82	8,86 (10,55%)	9994/10000
Politique RRP	0,008	11,78	7,87	0,005 (0,01%)	41/10000
$\Delta(\text{DRP},\text{RRP})$	-99,98%	-92,8%	-79,2%	-99,95%	-99,6%

Techniques	Violations V_{min} moyennes [$\text{m}^3/\text{scénario}$]	Pire scénario de violations cumulées [m ³]	Viol. instant. maximale [m ³]	Nb moy. de pdt avec violations [pas de tps/scén.]	Nbre de scénarios avec violations [scénarios]
Politique DRP	7,41	26,13	9,08	4,87 (5,8%)	9968/10000
Politique RRP	0,05	5,75	2,98	0,11 (0,13%)	991/10000
$\Delta(\text{DRP},\text{RRP})$	-99,32%	-78%	-67,2%	-97,75%	-90,06%

TAB. 4.7 – Déversés et violations de V_{min} pour RRP et DRP

Le résultat le plus remarquable dans Tab. 4.7 est la quasi-disparition des déversés et des violations moyens avec la politique RRP. Le pire scénario est celui pour lequel les violations cumulées sont les plus fortes. Nous voyons que RRP réduit ces violations par un facteur entre 4,5 et 15. Un autre indicateur est la violation maximale sur un pas de temps. La réduction est moins spectaculaire mais tout de même notable. Les deux dernières colonnes donnent la fréquence

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

des violations. RRP offre une sécurité remarquable. En résumé, RRP propose une gestion raisonnable des contraintes, alors que DRP n'évite des violations pour pratiquement aucun des scénarios, violations sans doute trop importantes en moyenne et en valeur maximale pour être acceptables.

Nous pouvons illustrer ces résultats à l'aide des cumulés de violations et de déversés :

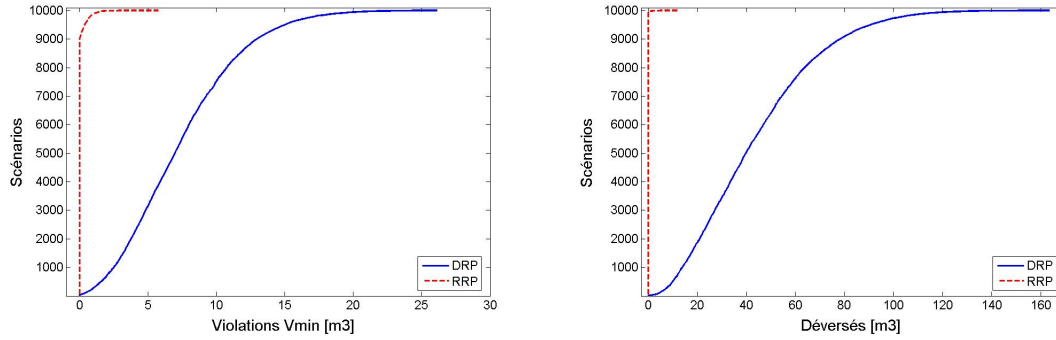


FIG. 4.7 – Répartition des violations cumulées par scénario sur V_{min} et V_{max} pour les politiques DRP et RRP.

En section 4.7.3, nous avons détaillé la façon dont les turbines prescrits par DRP et/ou RRP devaient être ajustés en temps réel pour réduire au maximum les violations de contraintes. Il est intéressant de considérer les valeurs moyennes et maximales des ajustements cumulés par scénario. Le tableau Tab. 4.8 donne ces valeurs. Nous pouvons voir que les ajustements pour RRP sont très faibles, ce qui rend crédibles les recommandations de turbines que RRP propose.

Ajustement cumulé [m^3]	DRP	RRP
Moyen	0,13	$8 \cdot 10^{-5}$
Maximal	2	$6 \cdot 10^{-1}$

TAB. 4.8 – Comparatif des ajustements infra-journaliers DRP et RRP

Pour conclure cette comparaison, le tableau Tab.4.10 récapitule les informations sur les coûts de IP et RRP, sans reprendre les informations sur RRP (puisque nous avons vu en section 4.8.1 que les coûts de DRP n'étaient pas pertinents).

Coût [Euros]	IP	RRP	$\Delta(\text{IP}, \text{RRP})$
Moyen	-1 092 493	-1 086 250	+ 0,57%
Min	-1 104 743	-1 098 454	+0,57%
Max	-1 079 167	-1 072 146	+ 0,65%

TAB. 4.9 – Coûts avec les politiques IP et RRP

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

Le tableau montre que RRP satisfait les contraintes avec un surcoût de seulement 0,57% par rapport au coût idéal en Information Parfaite. La figure 4.8 donne la distribution des coûts.

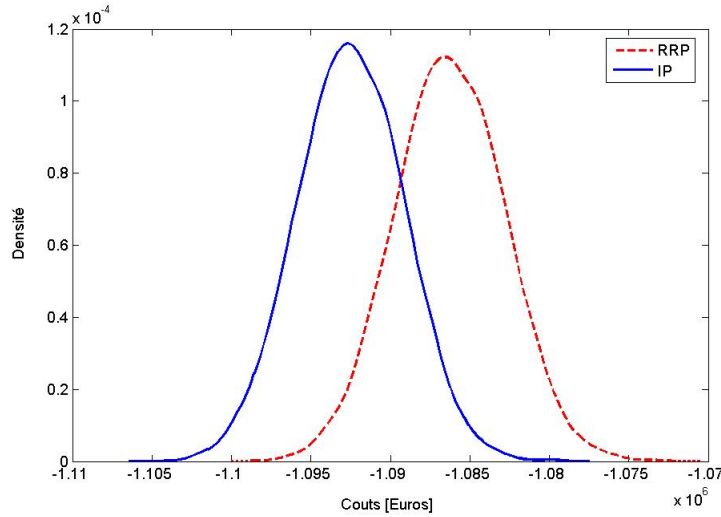


FIG. 4.8 – Densité de probabilité des coûts pour les modèles IP et RRP.

Un autre indicateur de prix est le prix moyen sur les scénarios sans aucune violation. Les prix sont ainsi représentatifs puisque le niveau de violation est identique. Les coûts du tableau suivant sont donc restreints aux 90% de scénarios pour lesquels les violations du modèle RRP sont nulles :

Coût [<i>Euros</i>]	IP	RRP	$\Delta(\text{IP}, \text{RRP})$
Moyen	-1 092 460	-1 086 218	+ 0,57%
Min	-1 104 742	-1 098 454	+0,57%
Max	-1 079 167	-1 072 146	+ 0,65%

TAB. 4.10 – Coûts avec les politiques IP et RRP sur les 90% de scénarios sans violation de contrainte

Nous constatons que les coûts moyens sont quasi-identiques que ceux du tableau 4.10. Cela signifie que :

- les coûts RRP sont significatifs malgré les violations restantes ;
- ces dernières violations sont si faibles qu'elles n'influent pas sur le coût.

L'écart de 0,57% entre IP et RRP est significatif et montre bien qu'une approche robuste ne détériore pas abusivement le coût moyen.

La dernière figure de cette section illustre les comportements d'une politique DRP et d'une politique RRP sur un scénario choisi au hasard (Fig.4.9). A première vue, les 2 modèles réagissent au signal de prix de manière assez semblable. Nous observons cependant que le modèle RRP se protège dans le réservoir de tête en turbinant moins, ce qui a pour effet, contrairement au modèle DRP, d'éviter les déversés de l'écluse intermédiaire. Par contre, dans le réservoir de

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

queue, le modèle RRP turbine davantage pour éviter d'atteindre la borne maximale de ce même réservoir. Ces écarts de volume expliquent la différence de coûts entre les 2 approches.

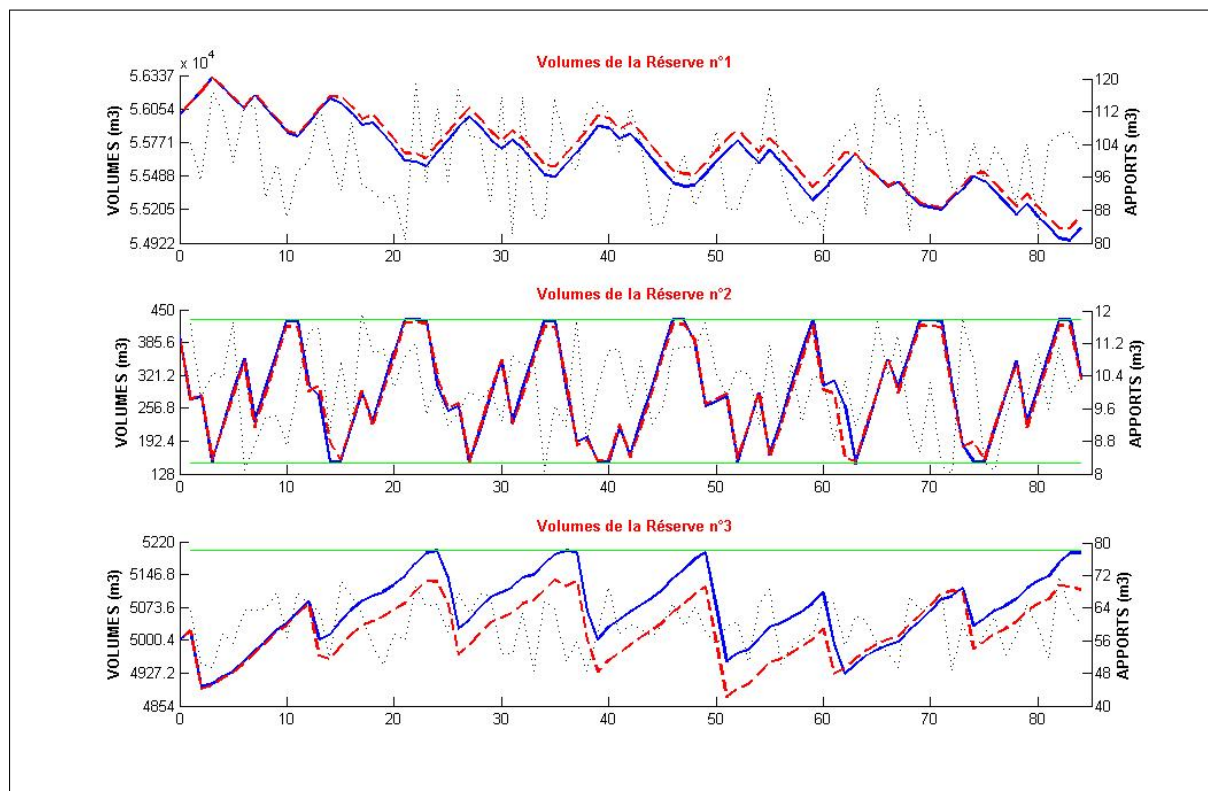


FIG. 4.9 – Comparaison des volumes après application des politiques DRP (en bleu) et RRP (en tirets rouges) sur un scénario d'apports donné (en pointillés).

Synthèse 4.2 (RRP vs. DRP)

Temps de calcul

Les modèles DRP et RRP ont des tailles comparables. Les temps de calcul sont donc très semblables.

Violations de contraintes

1. RRP annule les déversés moyens et divise le déversé maximum par près de 5. RRP rend le déversé exceptionnel : la fréquence est de 0,4% au lieu de 99,94% avec DRP.
2. L'amélioration de RRP sur DRP sur la gestion de la contrainte de volume minimum est comparable à celle sur le volume maximum.
3. Les turbines préconisés par l'approche RRP peuvent être mis en œuvre sans ajustement pour la quasi-totalité des scénarios.

Coûts

1. La très faible violation des contraintes assure la crédibilité au coût moyen de RRP (contrairement à DRP).

2. L'accroissement de coût avec RRP par rapport à une gestion en Information Parfaite n'est que 0,57%.

4.8.3 Comparaison entre RRP et RAL

Dans cette section, nous considérons un modèle avec règle d'ajustement linéaire (RAL) pour les décisions de turbinés à partir du deuxième jour. Nous présentons deux variantes. Le modèle RALD ne prend pas en compte les aléas sur les pas de temps de la première journée (comme présenté dans la section 4.5). Comme RRP, le modèle RALR prend en compte ces aléas par une approche robuste (sans ajustement linéaire sur la première journée). Quel que soit le modèle, nous n'introduisons que le dernier aléa dans la règle de décision (les justifications sont dans l'annexe A). Les résultats sont résumés dans le tableau Tab. 4.11.

Techniques	Déversé moy. [m ³ /scénario]	Pire scénario de déversés cumulés [m ³]	Déversé instant. maximal [m ³]	Nb moy. de pdt avec déversés [pas de tps/scén.]	Nbre de scén. avec déversés [scénarios]
<i>Politique DRP</i>	57,2	163,5	37,82	8,86 (10,55%)	9994/10000
<i>Politique RRP</i>	0,008	11,78	7,87	0,005 (0,01%)	41/10000
Politique RALD	3,08	65,74	29,77	1,005 (1,2%)	6299/10000
Politique RALR	1,52	65,74	29,77	0,285 (0,34%)	1829/10000

Techniques	Violations V_{min} moyennes [m ³ /scénario]	Pire scénario de violations cumulées [m ³]	Viol. instant. maximale [m ³]	Nb moy. de pdt avec violations [pas de tps/scén.]	Nbre de scénarios avec violations [scénarios]
<i>Politique DRP</i>	7,41	26,13	9,08	4,87 (5,8%)	9968/10000
<i>Politique RRP</i>	0,05	5,75	2,98	0,11 (0,13%)	991/10000
Politique RALD	0,96	34,30	14,9	0,59 (0,71%)	5449/10000
Politique RALR	0,09	17,22	9,66	0,06 (0,07%)	566/10000

TAB. 4.11 – Déversés et violations de V_{min} pour DRP, RRP, RALD et RALR

La première remarque est que les modèles RALD et RALR se comportent nettement mieux que DRP et presque aussi bien que RRP malgré l'absence de révision périodique. La deuxième remarque est que RALD et RALR ont des comportements presque identiques sauf sur le nombre de pas de temps et de scénarios avec violations. Il faut noter que ces violations dans le cas de RALD surviennent essentiellement en première journée, mais ces violations sont faibles et la protection assurée par l'approche robuste sur la première journée suffit à les éliminer. La dernière remarque est que les modèles RAL ont plus de difficultés à gérer les cas extrêmes.

La figure (Fig.4.10) représente la fonction de répartition des violations. Elle illustre le gain permis par RAL sur DRP et la proximité des résultats pour les politiques RAL et RRP. Nous observons enfin un comportement de RALR un peu meilleur que RALD, en particulier sur la contrainte de borne inférieure sur le volume.

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

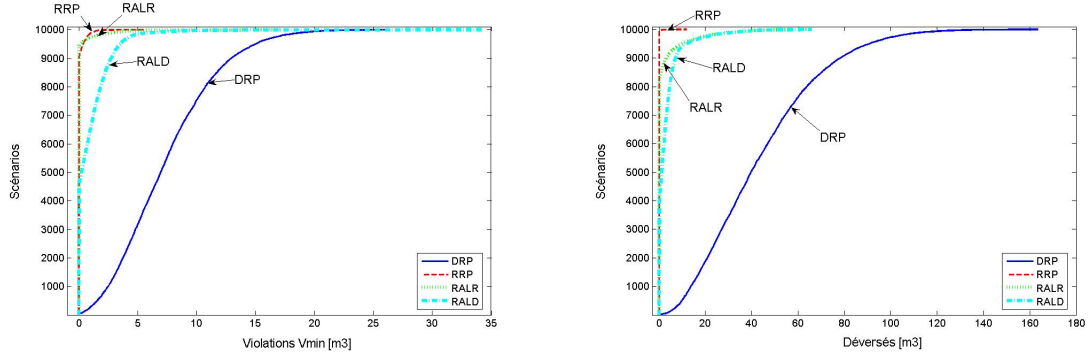


FIG. 4.10 – Violations cumulées par scénario sur V_{min} pour les modèles DRP, RRP, RALR et RALD.

Pour conclure sur les violations, les modèles RALR et RALD bénéficient eux aussi de peu d’ajustements infra-journaliers en moyenne. En revanche, et contrairement aux modèles à révisions périodiques, il peut arriver que sur certains scénarios, il faille faire des ajustements équivalents à 54 m^3 .

Ajustement cumulé [m^3]	RALR	RALD
Moyen	0,9	0,89
Maximal	53,66	53,67

TAB. 4.12 – Ajustements infra-journaliers RALR et RALD

Le dernier point à étudier est le prix de la couverture contre les aléas. Le tableau Tab.4.13 récapitule plusieurs informations intéressantes :

Coût [<i>Euros</i>]	IP	RRP	RALR	RALD
Moyen	-1 092 493	-1 086 250	-1 079 719	-1 079 952
Min	-1 104 743	-1 098 454	-1 088 679	-1 088 839
Max	-1 079 167	-1 072 146	-1 069 779	-1 070 113

TAB. 4.13 – Coûts avec IP, RRP, RALR et RALD

Les modèles RAL devant se couvrir sur du long terme sont plus conservatifs donc plus onéreux (pour un coût quasi similaire entre RALD et RALR). Cependant, le surcoût moyen de cette couverture n’est que de l’ordre de 1,2% par rapport au IP et de 0,6% par rapport au RRP. Cependant, les violations des modèles RAL peuvent biaiser légèrement ces coûts. Les courbes de distribution des coûts sont données à la figure 4.11.

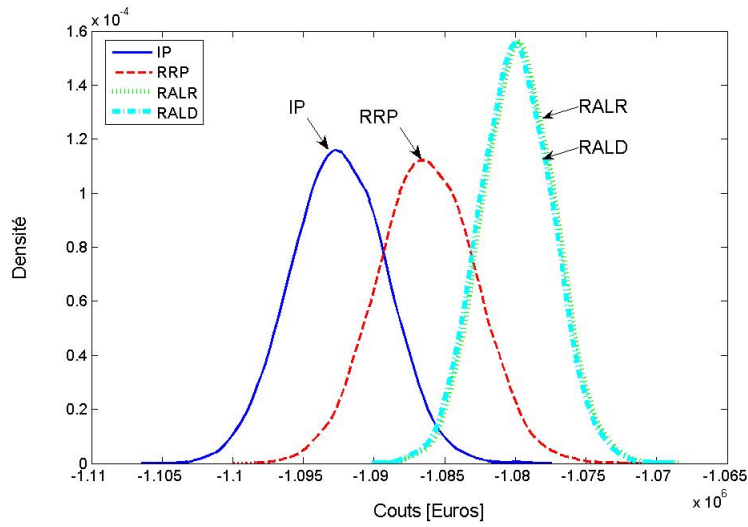


FIG. 4.11 – Densité des coûts pour les modèles IP, RRP, RALR et RALD.

Nous observons que les distributions RALD et RALR sont quasi-identiques et sont plus pointues et plus étroites que celles obtenues avec DRP et RRP, par effet de couverture. En effet, en se couvrant, le modèle cherche à être le plus proche d'une valeur moyenne pour assurer un coût moyen intéressant. Ceci explique le resserrement des densités et la réduction de l'écart-type.

La dernière figure (Fig.4.12) donne les profils de volume RRP et RALR sur un scénario particulier. La différence entre les 2 approches se fait essentiellement sur le long terme, où le modèle RALR montre une gestion plus précautionneuse. En effet, les ensembles d'incertitude en fin d'horizon sont plus grands qu'en premières périodes (puisque considérant plus d'aléas) rendant ainsi le "pire des cas" plus important.

4.8. EXPÉRIMENTATIONS NUMÉRIQUES

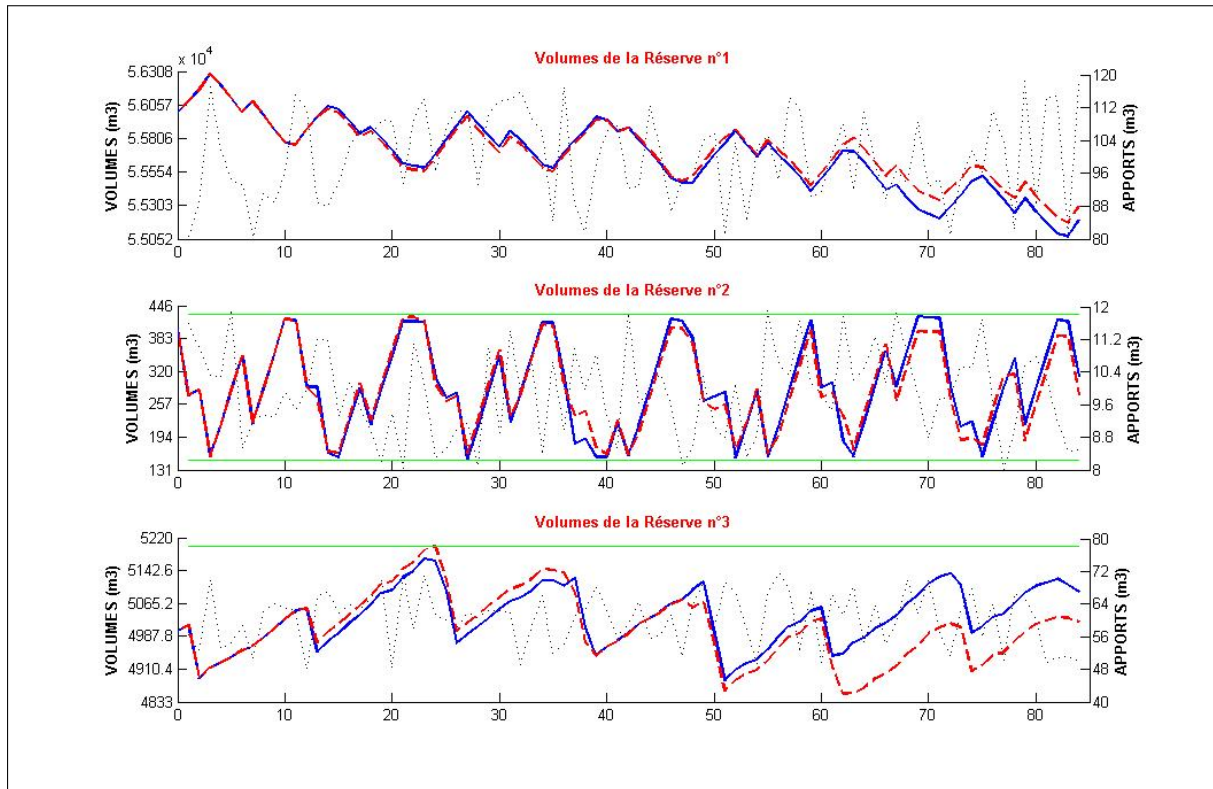


FIG. 4.12 – Evolution des volumes avec RRP (en bleu) et RALR (en tirets rouges) sur un scénario d'apports donné (en pointillés).

Synthèse 4.3 (RAL vs. DRP & RRP)

Politique de gestion

Comme DF, RAL fixe des règles de gestion qui ne sont pas révisées en fonction de l'information reçue.

Temps de calcul

Les modèles RAL sont beaucoup plus grands, mais ils restent accessibles aux codes d'optimisation linéaire (cf section 4.6). Ils sont trop coûteux en temps calcul pour être utilisés dans le cas de simulations à grande échelle³, mais ils peuvent très bien être utilisés dans une application en temps réel.

Violations de contraintes

1. RAL divise le déversé moyen de DF par 50 et celui de DRP par 30. La performance sur le déversé (en valeur instantanée ou cumulée) dans le pire scénario se situe entre les politiques en révision périodique DRP et RRP.
2. RAL a des performances comparables à RRP sur la gestion de la contrainte de volume minimum.
3. La mise en œuvre de RAL nécessite peu d'ajustements, sauf sur quelques scénarios exceptionnels.

Coûts

1. Le coût de gestion avec RAL n'est que de 0,58 % supérieur à celui de RRP et 1,15% à celui de IP. La faible valeur de la violation moyenne des contraintes donne de la crédibilité à la performance de coût observée.
2. Dans le cas du modèle simple avec 3 réservoirs, il n'y a pas de raison d'utiliser RAL. En revanche, sur des modèles plus complexes, la capacité d'anticipation de RAL peut s'avérer déterminante.

4.9 Expérimentations numériques : débit minimum

Nous testons le modèle robuste sur le même jeu d'essais mais en introduisant un débit minimum sur la vallée. Pour cela, nous obligeons des groupes à avoir un turbiné minimal. Les turbinés minimaux deviennent ainsi (en m^3/pdt) :

$$\underline{T}_g = \begin{pmatrix} 10 \\ 0 \\ 0 \\ 11 \\ 0 \\ 17 \end{pmatrix}$$

Nous récapitulons dans le tableau suivant tous les résultats :

Techniques	Déversé moy. [m^3 /scénario]	Pire scénario de déversés cumulés [m^3]	Déversé instant. maximal [m^3]	Nb moy. de pdt avec déversés [pas de tps/scén.]	Nbre de scén. avec déversés [scénarios]
Politique DRP	44,8	180,9	37,82	8,25 (9,8%)	9986/10000
Politique RRP	0,008	11,78	7,87	0,005 (0,01%)	42/10000
Politique RALD	2,78	65,74	29,77	0,71 (0,85%)	5831/10000
Politique RALR	1,48	65,74	29,77	0,283 (0,34%)	1809/10000

Techniques	Violations V_{min} moyennes [m^3 /scénario]	Pire scénario de violations cumulées [m^3]	Viol. instant. maximale [m^3]	Nb moy. de pdt avec violations [pas de tps/scén.]	Nbre de scénarios avec violations [scénarios]
Politique DRP	7,34	25,08	11,35	4,7 (5,6%)	9954/10000
Politique RRP	0,04	4,43	2,98	0,09 (0,1%)	823/10000
Politique RALD	0,88	22,9	7,1	0,57 (0,68%)	5376/10000
Politique RALR	0,078	19,53	7,1	0,06 (0,07%)	530/10000

TAB. 4.14 – Déversés et violations de V_{min} pour DRP, RRP, RALD et RALR

³Cette remarque doit être pondérée par le fait que nos expériences ont été réalisées avec un logiciel libre de modélisation et de résolution de programme linéaire. Les performances sont très inférieures à celles de logiciels commerciaux ou dédiés correspondants.

4.10. CONCLUSION

Les résultats sont assez proches que ceux obtenus sans le débit minimum. Les modèles robustes sont toujours aussi performants. Le modèle RALR est quasiment au niveau du RRP sans les révisions périodiques. Même les niveaux d'ajustement restent en faveur des modèles robustes (RRP) ou du même niveau que DRP :

Ajustement cumulé [m^3]	DRP	RRP	RALR	RALD
Moyen	0,13	2.10^{-4}	0,39	0,38
Maximal	2	8.10^{-1}	33,76	33,76

TAB. 4.15 – Ajustements infra-journaliers pour les modèles DRP, RRP, RALR et RALD

Enfin, nous pouvons comparer les coûts :

Coût [<i>Euros</i>]	IP	RRP	RALR	RALD
Moyen	-1 088 462	-1 082 171	-1 075 750	-1 075 750
Min	-1 101 298	-1 094 706	-1 084 547	-1 084 547
Max	-1 075 192	-1 067 931	-1 066 354	-1 066 354

TAB. 4.16 – Coûts avec IP, RRP, RALR et RALD

Nous retrouvons les écarts de coût faibles entre le coût idéal IP et les coûts robustes. Cet exemple supplémentaire montre bien que, quelles que soient les conditions initiales, les modèles robustes sont performants en terme de satisfaction des contraintes tout en assurant un coût moyen satisfaisant.

4.10 Conclusion

Le respect des contraintes sur le niveau des réservoirs est un point-clé de la gestion hydraulique avec apports d'eau aléatoires. Une approche classique fondée sur un modèle déterministe avec révision périodique ne permet pas d'éviter les violations de contraintes. Une approche par optimisation robuste, sans règle d'ajustement linéaires, mais avec révision périodique journalière, permet de respecter les contraintes avec un surcoût très faible par rapport au coût idéal d'une gestion en information parfaite.

Dans une approche à révision périodique, seules les décisions de première journée sont mises en œuvre. Les décisions sélectionnées pour les jours suivants peuvent être révisées. Elles ne sont pas contraignantes pour l'exploitant, mais leur calcul est nécessaire pour conduire de bonnes décisions le premier jour. Une question importante se pose. Etant donné que les décisions postérieures au premier jour peuvent tenir compte des réalisations des aléas du premier jour, faut-il incorporer cette information dans la modélisation des décisions ? Dans le modèle de la partie supérieure de la vallée de l'Ain, cela ne semble pas nécessaire. En effet, une projection déterministe semble suffire pour guider les décisions du premier jour, à condition que celles-ci soient prises dans le cadre d'un modèle robuste.

Si une anticipation plus fine s'avérait souhaitable, nous pouvons modéliser les décisions futures à l'aide de règles de décision linéaires. Nous l'avons fait sur un modèle à 7 jours et 84 pas de temps, sans révision périodique. La capacité d'anticipation de ce modèle est remarquable. Elle permet de contenir les violations de contrainte dans une limite raisonnable et cela avec un surcoût de seulement un demi pourcent du coût de la politique à révision périodique. La contrepartie d'un modèle à révision périodique est un accroissement substantiel (mais non exponentiel!) de la taille du modèle et du temps de calcul.

La présente étude tend à montrer que la programmation robuste est un outil adapté et efficace pour la gestion de modèles dynamiques complexes soumis à des incertitudes. Son atout majeur est le faible accroissement de taille de modèles linéaires lorsque nous passons du modèle déterministe à son équivalent robuste. Sa plus grosse limitation réside dans l'hypothèse que l'incertitude ne doit apparaître que dans le second membre des contraintes dynamiques. Cela exclut, dans un premier temps, une extension du modèle de gestion à des incertitudes sur les prix. D'autres approches, sans doute du type "programmation stochastique" comme dans [20], devraient être envisagées, éventuellement en les couplant avec une approche par programmation robuste pour la gestion d'aléas apparaissant dans le second membre.

Bien que nous ne l'ayons pas fait dans cette étude, il est tout à fait possible de définir une approche robuste incluant une modélisation fine des turbines de la première journée qui tiendrait compte des paliers de production à l'aide de variables entières. Cette modélisation ne peut être étendue aux règles de décisions linéaires qui par essence ne sont pas compatibles avec des restrictions entières. Cela n'est pas une restriction sévère, car, comme nous l'avons vu, les règles ne sont utilisées que pour modéliser les jours 2 et suivants et assurer une cohérence de plus long terme aux actions du premier jour.

Bibliographie

- [1] R. Apparigliato, J.-P. Vial, and R. Zorgati. Optimisation Robuste Linéaire - Application : Gestion Court Terme d'une Vallée Hydraulique. Technical report, EDF R&D, Dept OSIRIS, H-R32-2007-00658-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [2] R. Apparigliato, J.-P. Vial, and R. Zorgati. Gestion Hebdomadaire d'une Vallée Hydraulique par Optimisation Robuste. *Conférence conjointe FRANCORO-ROADEF, Grenoble (France)*, 23 Février 2007.
- [3] R. Apparigliato, J.-P. Vial, and R. Zorgati. Gestion Hebdomadaire d'une Vallée Hydraulique par Optimisation Robuste. *Séminaire scientifique de l'Institut Français du Pétrole, Rueil-Malmaison (France)*, 26 octobre 2007.
- [4] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. *19th International Symposium on Mathematical Programming, Rio de Janeiro (Brazil)*, August 2nd 2006.
- [5] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. *Third Romand Cycle of Operational Research, Zinal (Switzerland)*, March 5th 2007.

- [6] R. Apparigliato, J.-P. Vial, and R. Zorgati. An Original Approach for the Weekly Management of a Hydraulic Valley. *Third Romand Cycle of Operational Research, Zinal (Switzerland)*, March 7th 2006.
- [7] R. Apparigliato, J.-P. Vial, and R. Zorgati. Weekly Management of a Hydraulic Valley by Robust Optimization. In F. S. Hillier, editor, *Handbook on Uncertainty and Environmental Decision Making*, International Series in Operations Research and Management Science. Springer, to appear (submitted).
- [8] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-Ph. Vial. Retailer-Supplier Flexible Commitments Contracts : a Robust Optimization Approach. *Manufacturing and Service Operations Management*, 7 :248–271, 2005.
- [9] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming*, 99(2) :351–376, 2004.
- [10] D. Dentcheva and W. Römisch. Optimal Power Generation under Uncertainty via Stochastic Programming. *Stochastic Programming Methods and Technical Applications*, 458 :22–56, 1998.
- [11] S.-E. Fleten and T.K. Kristoffersen. Short-Term Hydropower Production Planning by Stochastic Programming , 2008.
- [12] N. Gröwe-Kuska, K.C. Kiwiel, M.P. Nowak, W. Römisch, and I. Wegner. Power Management in a Hydro-Thermal System under Uncertainty by Lagrangian Relaxation. In A. Borghetti, C.A. Nucci, and M. Paolone, editors, *IMA Volumes in Mathematics and its Applications*, volume 128, pages 39–70. Springer-Verlag, New-York, 2002.
- [13] N. Gröwe-Kuska and W. Römisch. *Stochastic Unit Commitment in Hydro-Thermal Power Production Planning*, chapter 30. MPS-SIAM Series in Optimization, 2005.
- [14] H. Heitsch and W. Römisch. Generation of Multivariate Scenario Trees to Model Stochasticity in Power Management. *Computational Optimization and Applications*, 24(2-3) :187 – 206, February-March 2003.
- [15] B. Mo, A. Gjelsvik, and A. Grundt. Integrated Risk Management of Hydro Power Scheduling and Contract Management. *IEEE Transactions on Power Systems*,, 16(2) :216–221, 2001.
- [16] R. Nürnberg and W. Römisch. A Two-Stage Planning Model for Power Scheduling in a Hydro-Thermal System under Uncertainty. *Optimization and Engineering*, 3 :355–378, 2002.
- [17] M. V. F. Pereira and L. M. V. G. Pinto. Multi-Stage Stochastic Optimization Applied to Energy Planning. *Mathematical Programming*, 52 :359–375, 1991.
- [18] A. Prékopa and T. Szántai. On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake. *European Journal of Operations Research*, 3 :175–189, 1979.
- [19] T.A. Rotting and A. Gjelsvik. Stochastic Dual Dynamic Programming for Seasonal Scheduling in the Norwegian Power System. *IEEE Transactions on Power Systems*,, 7(1) :273–279, 1992.
- [20] J. Thénie and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : a Heuristic Approach. *Automatica*, 2008 (à paraître).

BIBLIOGRAPHIE

- [21] G. Unger. *Hedging Strategy and Electricity Contract Engineering*. PhD thesis, ETH, Zurich, 2002.
- [22] G. Vignal. Modèle APOGEE. Vers une Nouvelle Formulation des Sous-Problèmes Hydrauliques. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/003/A, 2004.
- [23] R. Zorgati and P. Thomas. Formation au Modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d'une Vallée sur Critère Economique. I. Formulation. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/015/A, 2004.
- [24] R. Zorgati and P. Thomas. Formation au modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d'une Vallée sur Critère Economique. II. Résolution. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/016/A, 2004.

Troisième partie

Gestion de la couverture contre le risque de défaillance physique

Chapitre 5

La gestion du risque physique

Sommaire

5.1	Introduction à la gestion du risque physique	152
5.1.1	Obligation réglementaire de gestion du risque de défaillance physique . .	152
5.1.2	Procédure actuelle de couverture, limitations	152
5.1.3	Les outils financiers disponibles sur les marchés	153
5.2	Etude de l'existant à EDF	154
5.3	Une approche opérationnelle simplifiée	159
5.4	Modélisation du problème en boucle ouverte	163
5.4.1	Boucle ouverte et boucle fermée	163
5.4.2	Formulation du problème avec les actifs financiers	164
5.4.3	Calcul des quantiles	170
5.4.4	Résultats	170
5.5	Conclusion	171
	Bibliographie	171

La couverture contre le risque physique à l’horizon hebdomadaire n’a été jusqu’ici que peu étudiée. Aucun réel outil d’optimisation n’est à disposition des exploitants, mis à part leurs propres heuristiques. Dans ce chapitre, nous formulons le problème de couverture comme un modèle d’optimisation stochastique dans lequel la marge est vue comme un processus stochastique indépendant des actions de couverture. Nous proposons également des méthodes de résolution à ce problème et analysons les solutions. La procédure de ce chapitre a été implémentée et donne lieu à une maquette (en langage Matlab avec appel du kit libre de modélisation GLPK pour la résolution des problèmes d’optimisation linéaires) pouvant servir de démonstration aux utilisateurs. Les travaux de ce chapitre ont donné lieu à une note interne [3] et à des présentations orales lors de congrès internationaux ([2],[1]).

5.1 Introduction à la gestion du risque physique

5.1.1 Obligation réglementaire de gestion du risque de défaillance physique

L’électricité étant, par nature, difficile à stocker, elle doit être disponible à chaque instant pour satisfaire la consommation et ainsi assurer l’équilibre entre l’offre et la demande. Nous définissons pour cela la marge de production comme l’écart entre l’offre totale et la demande totale en électricité, compte tenu des aléas affectant le système électrique.

Compte tenu des dispositions réglementaires en matière de gestion du risque de défaillance physique, l’objectif de chaque responsable d’équilibre est d’assurer dans son périmètre l’équilibre offre-demande dans 99% des situations, avec un recours aux marchés de l’électricité réduit à ses engagements antérieurs à moyen-terme. Cette politique se traduit dans la pratique par la définition d’une marge de sécurité, dite “marge au risque 1%”, qui représente la marge (exprimée en MW) pour que la probabilité d’avoir une marge de production M négative soit égale à $1 - p\%$ ($p = 1\%$), soit $Prob(M < 0) \geq 1 - p$.

5.1.2 Procédure actuelle de couverture, limitations

Lorsque le risque de défaillance physique est trop élevé, l’exploitant cherche à se couvrir par des décisions de couverture telles que : le déclenchement d’options tarifaires (contrats d’effacement), des achats de sécurité sur les marchés ou, plus rarement, l’appel à des unités de production de pointe voire la modification de la composition du parc de production (typiquement un départ en maintenance ou en essai d’une centrale). D’ailleurs, l’objectif prioritaire étant le retour à l’équilibre offre-demande, nous pouvons être amenés à payer très cher cette couverture.

Le processus actuel de gestion du risque à l’horizon court-terme se caractérise par trois points majeurs :

- le calcul de la marge de production est effectué à partir du planning de production, fixé et obtenu par optimisation déterministe (le planning de maintenance est déterminé) ;
- la marge au risque 1% est évaluée selon un calcul analytique issu du “moyen terme” et adapté de façon ad hoc à l’horizon hebdomadaire. Ce calcul est fondé sur l’hypothèse que les aléas

affectant la demande de consommation, l'hydraulicité et la disponibilité du parc sont des processus aléatoires gaussiens ;

- le processus de couverture est manuel : si le risque de défaillance physique est jugé trop important, il est décidé à dire d'expert quels produits de couverture utiliser (options tarifaires, achats de sécurité sur les marchés et en dernier recours, la modification du planning de maintenance de groupes thermiques).

Les limites d'un tel processus sont :

- la non-adéquation à l'horizon court-terme de l'hypothèse de normalité des processus aléatoires dans le calcul de la marge au risque 1%.
- la non-optimalité économique des décisions de couverture ;
- éventuellement, la non prise en compte du risque financier.

5.1.3 Les outils financiers disponibles sur les marchés

Il existe actuellement sur les différents marchés de l'électricité plusieurs types de contrats financiers disponibles. Les marchés qui nous concernent principalement sont les marchés français Powernext (<http://www.powernext.fr>) et allemand EEX (<http://www.eex.com>). Des descriptifs plus précis sont disponibles sur les sites web de ces bourses.

Le Spot

Le Spot est le marché du jour pour le lendemain. Les membres peuvent négocier des contrats horaires standardisés qui les engagent à acheter ou à vendre des volumes d'électricité, pour une heure donnée, à un prix de marché sur l'ensemble du réseau. Plusieurs heures peuvent également être liées en blocs. Ces contrats sont des contrats à terme ferme sur marchandise dont le sous-jacent est l'électricité. Nous considérerons deux types de contrats : Spot Base et Spot Pointe en fonction du type de blocs horaires souhaités.

Le Week Ahead

Les contrats Week Ahead sont des produits à terme qui peuvent être achetés (ou vendus) à chaque jour ouvrable pour une livraison la semaine suivante. Les quantités livrées sont égales pour chaque jour ouvrable de la semaine $S + 1$. Tout comme pour le marché Spot, il existe différents blocs correspondant aux heures de livraison. Nous distinguerons les Week Ahead Base et les Week Ahead Pointe.

Le Week-End Ahead

Les contrats Week End-Ahead sont des produits à terme qui peuvent être achetés (ou vendus) à chaque jour ouvrable pour une livraison le week-end suivant. Les quantités livrées le samedi et le dimanche sont égales. Tout comme pour le marché Spot, il existe différents blocs correspondants

aux heures de livraison. Nous distinguerons les Week End-Ahead Base et les Week End-Ahead Pointe.

Les options tarifaires

Les options tarifaires sont des contrats spécifiques à EDF, principalement pour les gros clients. Il en existe plusieurs types. Celles que nous considérerons dans cette étude sont les EJP (Effacements Jours de Pointe). Les clients disposant d'un tel type de contrat bénéficient d'un tarif préférentiel durant toute l'année (le prix du KWh est proche de celui des heures creuses). Cependant, EDF dispose de 22 jours pour lesquels le prix de l'électricité est si élevé que le client est incité à délester sa consommation. Ce sont donc 22 jours d'effacement de production qui constituent un stock d'électricité pour EDF. Les clients sont prévenus la veille pour le lendemain.

Cette option n'est plus proposée et a été remplacée pour les nouveaux clients par l'option TEMPO. Le principe est le même avec une offre plus flexible (introduction de jours bleus, blancs, rouges et l'introduction d'heures pleines et creuses).

Pour simplifier, nous appellerons options tarifaires un stock de jours dont la production est effaçable entièrement, à un prix contractualisé avec une décision à prendre la veille pour le lendemain.

5.2 Etude de l'existant à EDF

Comme nous l'avons présenté au chapitre 1, la gestion de la production à l'horizon hebdomadaire joue deux rôles principaux : la sentinelle su journalier et la gestion du risque physique. L'approche naturelle qui paraît optimale est de résoudre conjointement les problèmes déduits de ces deux rôles. Les décisions de placement du parc de production et d'achats/ventes sur les marchés seraient ainsi optimales économiquement tout en satisfaisant, d'une part, les contraintes techniques et, d'autre part, le risque de défaillance physique. Cependant, un tel modèle peut être difficile à mettre en place, affirmation que nous allons illustrer dans la suite de cette section.

Le problème de la gestion de la production consiste, à l'horizon hebdomadaire, à trouver un planning de production :

- optimal économiquement,
- satisfaisant l'ensemble des contraintes techniques de chaque unité de production,
- satisfaisant l'équilibre offre-demande à chaque instant.

Nous pouvons formuler ce problème de gestion de production sous la forme générale :

$$\left\{ \begin{array}{l} \min_{\mathbf{P}} \sum_{i=1}^N \sum_{t=1}^T C_{i,t}(P_{i,t}) \\ \text{sous} \\ \sum_{i=1}^N P_{i,t} = D_t, \forall t \in 1, \dots, T \\ P_{i,t} \in U_{i,t}^{ad}, \forall i \in 1, \dots, N, \forall t \in 1, \dots, T, \end{array} \right. \quad (5.1)$$

où N est le nombre d'unités de production, T le nombre de pas de temps, \mathbf{P} est la matrice $N \times T$ contenant toutes les productions $P_{i,t}$ produites par l'unité de production i à l'instant t . $C_{i,t}(P_{i,t})$ est le coût de production de i au pas de temps t , $U_{i,t}^{ad}$ son ensemble de réalisabilité (l'ensemble des contraintes techniques) et D_t la demande en électricité à satisfaire en t .

Ce problème s'avère très difficile puisqu'il est :

- **de grande taille** : le problème doit prendre en compte 58 tranches nucléaires, une centaine de centrales hydrauliques réparties sur 40 vallées, une cinquantaine de centrales thermiques et tous les différents types de contrats d'effacement (EJP, Tempo,...) passés avec des clients. D'autre part, chaque unité a ses propres contraintes techniques à satisfaire ;
- **non linéaire** : en raison des coûts de démarrages et de certaines contraintes techniques non linéaires (rendement des turbines, hauteur de chute,...) ;
- **non convexe** : notamment à cause des paliers de production, des décisions marche/arrêt du thermique ou des coûts de démarrage ;
- **à états discrets** : en particulier à cause des paliers de production, des décisions binaires d'effacement,...

De plus, la contrainte de satisfaction de la demande en électricité à chaque instant couple toutes les unités de production (quel que soit leur type). Sans cette contrainte, le problème est trivial puisqu'il est possible de le décomposer en problèmes indépendants simples, dont la solution minimale est une production nulle. D'autres contraintes couplantes du même type, liées à la gestion des réserves, sont également ajoutées au problème.

Pour résoudre ce problème à l'horizon hebdomadaire, soit sur 12 jours avec un pas de temps bi-horaire (donc 144 pas de temps), EDF dispose d'un modèle ([17],[16],[18],[19],[20],[15]) basé sur une décomposition par les prix (algorithme de décomposition-coordination) permettant une gestion séparée des différentes unités de production. Chaque unité de production est gérée indépendamment et doit résoudre son problème d'optimisation de la production compte tenu de son coût propre de production par rapport à la rémunération fournie par la coordination (résolue à l'aide d'une méthode de faisceaux), sous ses contraintes locales. Chaque sous-problème :

- thermique est résolu par programmation dynamique. Un graphe d'états permet de modéliser les différents paliers de production ainsi que les contraintes de fonctionnement comme la durée

minimale sur un palier ;

- hydraulique (relatif à une vallée hydraulique) est un problème linéaire en phase de Lagrangien simple et un problème quadratique en phase de Lagrangien augmenté (voir chapitre 4) résolu par une méthode de points intérieurs.

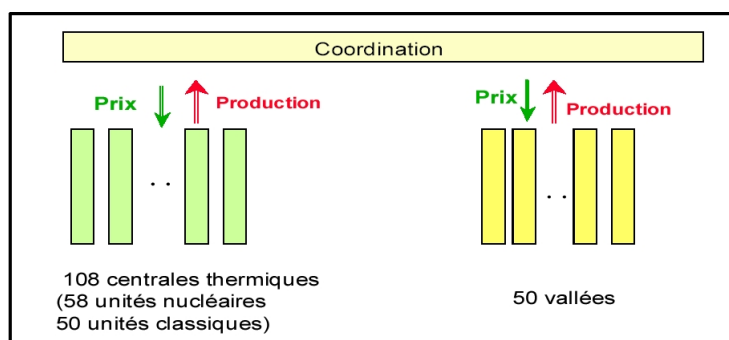


FIG. 5.1 – Schéma de la décomposition par les prix.

La principale caractéristique de ce modèle est sa vision déterministe. Au court terme, nous ne considérons pour le moment qu'une vision moyenne de ces aléas et supposons donc que les prévisions sont suffisamment pertinentes. Cet outil permet de résoudre le problème de production avec une modélisation fine du fonctionnement des unités de production, en des temps de calculs relativement réduits (de l'ordre d'une dizaine de minutes pour l'ensemble du parc). Pourtant, même à un horizon court terme, quatre aléas¹ jouent un rôle déterminant à partir de quelques jours :

- la **demande en électricité**, fortement corrélée à la température ;
- l'**hydraulicité** dépendant de nombreux phénomènes climatiques comme les précipitations, la fonte des neiges,... ;
- l'**indisponibilité fortuite** des différents types de centrales ;
- les **prix** : de marché, des matières premières,...

Les décisions doivent être prises en tenant compte des réalisations des aléas affectant le système. En effet, une prise en compte insuffisante de ces processus aléatoires peut engendrer une mauvaise gestion de l'équilibre offre-demande ou une mauvaise anticipation des commandes futures.

Le problème de planification (5.1) peut être étendue sous sa forme générale au cas stochastique :

¹Dire que quatre aléas principaux influent sur le système électrique est un réel abus de langage. L'aléa de prix est un aléa multi-dimensionnel. En effet, si nous considérons un marché Spot et un marché à terme, nous disposons de deux aléas de prix différents (éventuellement corrélés). La complexité s'accroît ainsi très rapidement.

$$\left\{ \begin{array}{l} \min_{\mathbf{P}} \mathbb{E} \left\{ \sum_{t=1}^T \left[\sum_{i \in I^{Th}} C_{i,t}(P_{i,t}^{Th}(\tilde{\xi}_{[t]}^I), \tilde{\xi}_{[t]}^P) + \sum_{i \in I^{Hyd}} C_{i,t}(P_{i,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H), \tilde{\xi}_{[t]}^P) \right] \right\} \\ \text{sous} \\ \sum_{i \in I^{Hyd}} P_{i,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H) + \sum_{i \in I^{Th}} P_{i,t}^{Th}(\tilde{\xi}_{[t]}^I) = \tilde{D}_t, \forall t \in 1, \dots, T \\ P_{i,t}^{Th}(\tilde{\xi}_{[t]}^I) \in U_{i,t}^{Th-ad}, \forall i \in I^{Th}, \forall t \in 1, \dots, T \\ P_{i,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H) \in U_{i,t}^{Hyd-ad}, \forall i \in I^{Hyd}, \forall t \in 1, \dots, T, \end{array} \right. \quad (5.2)$$

où I^{Th} est l'ensemble des centrales thermiques, nucléaires,... et I^{Hyd} l'ensemble des centrales hydrauliques. Nous avons évidemment $|I^{Th}| + |I^{Hyd}| = N$. Dans le problème (5.2), les quatre processus stochastiques pris en compte sont la demande en électricité \tilde{D}_t , les prix $\tilde{\xi}^P$, l'hydraulicité $\tilde{\xi}^H$ et l'indisponibilité des centrales $\tilde{\xi}^I$. Pour chacun de ces aléas, nous introduisons la notation $\tilde{\xi}_{[t]} = (\xi_1, \dots, \xi_t)$ pour décrire son histoire jusqu'à t . Les productions des centrales thermiques et nucléaires agrégées en $P_{i,t}^{Th}$ peuvent dépendre de l'aléa d'indisponibilité et doivent satisfaire les contraintes techniques $U_{i,t}^{Th-ad}$. Les productions des centrales hydrauliques sont, elles, fonction des aléas d'indisponibilité et d'hydraulicité et doivent satisfaire les contraintes techniques $U_{i,t}^{Hyd-ad}$. Les coûts sont dépendants des aléas sous-jacents aux productions ainsi qu'aux prix. La complexité du problème est accrue en raison des multiples aléas. Les productions dépendent de l'aléa, mais uniquement de l'histoire (ici jusqu'à t , en supposant que les décisions sont prises en connaissance de la réalisation de l'aléa du jour) pour ne pas violer le principe de non-anticipativité. La contrainte de satisfaction de la demande est ici écrite sous forme d'égalité mais il s'agit d'un abus d'écriture. En effet, sauf dégénérescence, une telle contrainte ne peut pas être satisfaite, sauf si elle contient une variable d'ajustement automatique fixée après la réalisation de l'aléa.

Pour appréhender cette nouvelle difficulté, différentes approches ont été testées. Une approche multi-scénarios [21] consistant à lancer plusieurs fois le modèle déterministe précédent sur différents scénarios (extrêmes et moyen) a été mise en place. Elle permet de borner les coûts et de fournir ainsi des indicateurs très utiles à l'exploitant. Une autre approche a été mise en place mais, pour le moment, à un horizon moyen-terme (1 an). Il s'agit de résoudre le même problème sur un arbre d'aléas avec une démarche de programmation stochastique ([13],[14]). Comme cet outil est prévu pour le moyen-terme, la description du parc est bien moins détaillée : toutes les vallées hydrauliques sont agrégées, les contraintes techniques non convexes ont été en partie supprimées ou simplifiées tout comme les contraintes dynamiques,... Les temps de calculs de ce modèle sont de plusieurs minutes avec au final un écart de bouclage de quelques centaines de MW. Si nous décrivions encore plus finement le parc de production, avec en particulier l'introduction de variables binaires et de contraintes non convexes, le problème serait impossible à résoudre sur les 144 périodes de temps de la période d'étude dans des temps compatibles avec une utilisation industrielle.

La prise en compte d'aléas dans la gestion de production hebdomadaire est donc déjà un problème d'une difficulté majeure. Mais pour satisfaire les exigences que nous nous sommes fixées, le dernier point à intégrer dans le problème est la politique de risque du producteur EDF. Nous avons vu en section 5.1.1 que le critère appliqué en politique de risque est le critère 1%, que

Mais dans le cas général, ce type de contrainte est source de grandes difficultés, les contraintes en probabilité définissant un ensemble de réalisabilité généralement non convexe voire non connexe (chap. 3 et nous pouvons citer notamment [11],[10],[12],[6],[5]). D'ailleurs, à l'heure actuelle, seuls les cas de lois logarithmiquement quasi-concaves ou elliptiques symétriques peuvent assurer la convexité, conditions qui ne sont a priori pas satisfaites par les aléas du problème. Nous pouvons remarquer que la contrainte de demande, désormais satisfaite en probabilité, est devenue une inégalité. Comme nous en avons déjà discuté, l'égalité ne peut pas être satisfaite en présence d'aléas. Nous introduisons donc implicitement un ajustement qui est la violation, ajustement non souhaité, mais non comptabilisé par un quelconque coût de défaillance.

L'idée d'un outil global est donc quelque chose d'imaginable à moyen terme mais son élaboration nécessite encore des années de travail de recherche, notamment pour ce qui concerne des sujets tels que la modélisation et la prise en compte des aléas, les problèmes non convexes de grande taille, les contraintes en probabilité... Afin de répondre aux besoins industriels à court terme, nous proposons, à l'image du processus métier, de dissocier la gestion de la production hebdomadaire et la gestion de la couverture physique.

5.3 Une approche opérationnelle simplifiée

Le problème (5.3) est donc un problème très difficile à résoudre. Une possibilité, motivée par les pratiques métier actuelles, est la dissociation de la gestion de la production et de l'optimisation de la couverture. Cette approche, bien que sous-optimale, a le mérite de simplifier le problème et constitue donc une alternative intéressante. Cependant, pour que cette séparation soit possible, nous devons introduire un indicateur de l'état de parc de production. En effet, les décisions de couverture doivent être prises en connaissance du planning de production, notamment pour que les interventions sur les marchés soient optimales. C'est pour cela que nous définissons la notion fondamentale de **marge de production** :

Définition 5.1 *La marge prévisionnelle de production (exprimée en MW) à l'instant t est l'écart entre la capacité maximale de production électrique du parc (sans marché et options tarifaires) et la demande en électricité à t . Formellement, la marge peut s'écrire :*

$$\tilde{M}_t = C_{\max,t}^{Th}(\tilde{\xi}_{[t]}^I) + C_{\max,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H) - \tilde{D}_t,$$

où :

- $C_{\max,t}^{Th}(\tilde{\xi}_{[t]}^I)$ est la capacité maximale de production à l'aide de moyens thermiques,
- $C_{\max,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H)$ est la capacité maximale de production hydraulique,
- \tilde{D}_t est la demande en électricité à l'instant t .

Etant dépendante de paramètres aléatoires, la marge est donc un processus stochastique.

Elle dépend de l'histoire des processus puisque, par exemple, la capacité maximale hydraulique à t dépend de tous les apports depuis le début de la période.

Le terme $\Lambda_t = \sum_{i \in I^{Hyd}} P_{i,t}^{Hyd}(\tilde{\xi}_{[t]}^I, \tilde{\xi}_{[t]}^H) + \sum_{i \in I^{Th}} P_{i,t}^{Th}(\tilde{\xi}_{[t]}^I) - \tilde{D}_t$ dans la contrainte en probabilité représente la différence entre la production planifiée et la demande. Si nous conservons cette

écriture, nous allons prendre des décisions d'achats et de ventes basée sur la vision d'un parc figé en déterministe. Or, une perturbation δ_t sur les prévisions d'aléas interviendra forcément. Dans le cas, par exemple, de pannes de centrales ou d'une augmentation soudaine de la température, les recours X_t prévus deviennent alors insuffisants. En effet, dans les deux cas, Λ_t devient moins important puisque soit la production a baissé (cas de la panne), soit la demande a augmenté (cas de la hausse de température). Il faut donc avoir un recours $X_t(\delta_t)$ supplémentaire qui peut être payé à un prix qui peut avoir lui aussi augmenté. Nous sentons bien que se baser sur un parc figé en vision déterministe est particulièrement risqué physiquement et financièrement. Pour réduire cette part de risque, nous pouvons ainsi estimer la capacité maximale d'électricité que le parc de production en l'état est capable de fournir et la substituer à la production planifiée. Comme cela :

- nous introduisons une marge de manoeuvre en cas de réalisation d'aléa défavorable ;
- nous limitons également les appels aux marchés et nous favorisons ainsi l'appel au parc de production (même si les appels au parc ne sont pas optimisés car non pris en compte en tant que variables de décision dans le modèle).

Mais, finalement, nous venons de redéfinir la marge de production.

La procédure d'obtention de cette marge est présentée dans le chapitre 7. Dans cette section, nous supposons donc que la marge de production est une donnée et est disponible sous forme de scénarios. Nous pouvons tout de même rappeler les hypothèses de base du calcul de marge :

- la marge de production nécessite de connaître l'état du parc à l'instant initial d'étude. Cet état initial est la solution du problème déterministe de gestion de production hebdomadaire. La phase d'optimisation du parc de production demeure donc primordiale, même si elle est désormais dissociée de l'optimisation de la couverture ;
- pour des raisons de temps de calcul, la marge de production n'est plus calculée par pas de temps mais par jour. Nous ne calculons plus qu'une marge par jour : la marge à l'heure de pointe. Cette simplification se justifie par le constat que si la quantité achetée est suffisante pour satisfaire l'heure de pointe, alors elle sera suffisante pour passer toute heure de la journée. Cependant, la couverture finale conduit à une sur-protection.

Nous pouvons donc énoncer l'hypothèse de base de notre approche :

Hypothèse 5.1 *Nous choisissons de découpler la gestion de production et le processus de couverture. La gestion de la production à court terme est toujours effectuée en déterministe à l'aide des modèles disponibles. L'optimisation des décisions de couverture contre le risque physique est stochastique et est basée sur la volonté de s'assurer, pour chaque jour, d'une marge de production à la pointe supérieure à une marge de sécurité fixée, sur au moins $(100-p)\%$ des scénarios disponibles.*

Le problème de couverture contre le risque physique s'écrit dorénavant :

$$\left\{ \begin{array}{l} \min_{\mathbf{X}_j} \mathbb{E} \left\{ \sum_{j=1}^J \sum_{i \in I^X} C_j^i(X_j^i, \xi_{[j]}^{\tilde{P}}) \right\} \\ \text{sous} \\ Prob(\tilde{M}_j + \sum_{i \in I^X} X_j^i \geq M_S) \geq 1 - p, \forall j \in 1, \dots, J \\ \tilde{M}_j = C_{\max,j}^{Th}(\xi_{[j]}^I) + C_{\max,j}^{Hyd}(\xi_{[j]}^I, \xi_{[j]}^H) - \tilde{D}_j, \forall j \in 1, \dots, J \\ X_j^i \in U_{i,j}^{X-ad}, \forall i \in I^X, \forall j \in 1, \dots, J, \end{array} \right. \quad (5.4)$$

où les X_j^i sont les décisions de sécurité déjà définies dans la section précédente et $U_{i,j}^{X-ad}$ les contraintes techniques correspondantes. La contrainte en probabilité introduite peut être appa-
rentée à l'obligation de service public.

Commentaires sur l'approche : nous pouvons brièvement commenter quelques conséquences du choix de cette formulation du problème de couverture :

- **la pratique des exploitants** : cette approche est principalement calquée sur le mode opératoire des exploitants (cf. sections 5.1.1 et 5.1.2). Historiquement, chaque producteur est tenu d'assurer l'équilibre offre-demande. Quand les marchés n'existaient pas encore, il était nécessaire de s'assurer que la capacité maximale de production disponible sur la période d'étude était suffisante pour satisfaire la demande tout en pouvant faire face aux survenues de pannes pouvant affecter le système. Comme les producteurs continuent de maintenir l'équilibre offre-demande et que les quantités d'électricité disponibles à l'heure actuelle sur les marchés ne sont pas encore suffisamment importantes (en particulier à cause des faibles capacités d'interconnexions du réseau entre pays), un tel processus conserve encore tout son intérêt. Nous pouvons toutefois souligner que cette approche améliore la pratique des exploitants : nous avons réussi à formuler leur pratique pour pouvoir le résoudre à l'aide de solveurs adaptés. De plus, nous avons rendu le calcul de marge plus réaliste en ne considérant plus des aléas gaussiens (voir chapitre 7) ;
- **la simplification algorithmique du problème** : les simplifications apportées sont importantes. Le modèle (5.3) comporte des contraintes techniques sur les unités de production physique (paliers de production, courbes de démarrage, d'arrêt,...). Ces contraintes, en général non convexes, ne figurent plus dans le modèle (5.4), conduisant à des solutions sur-optimales (car moins contraintes). C'est la contrepartie d'une plus grande facilité de résolution numérique ;
- **la réduction de la taille du problème** : nous ne considérons que le respect de la marge à l'heure de pointe de chaque jour de la période de temps. Nous passons ainsi de 144 pas de temps à uniquement 12. Par cette hypothèse, nous considérons que, si nous sommes capables de satisfaire la demande à la pointe, nous serons capables de satisfaire la demande à tout autre instant de la journée. La cause principale de ce choix de modélisation provient du calcul de la capacité de production maximale hydraulique. Le chapitre 7 montre que, pour calculer cette quantité, il est nécessaire de résoudre à chaque instant autant de problèmes linéaires que de vallées hydrauliques (une quarantaine) pour chaque scénario d'hydraulicité. Nous voyons

aisément que les temps de calculs peuvent devenir excessifs quand le nombre de pas de temps devient trop important. Nous réduisons donc ce nombre de pas de temps. Là encore, cette hypothèse est forte mais cohérente avec les pratiques des exploitants. En effectuant cette hypothèse, nous limitons les arbitrages sur les marchés dans la journée (chaque journée est réduite à son heure de pointe). Mais comme des produits horaires ne sont pas encore trop développés, les arbitrages sont rares et les conséquences sur la couverture sont donc limitées ;

- la **sous-optimalité et la sur-couverture du problème** : comme les processus de décision sont dissociés, il est évident que les décisions prises sont sous-optimales. Deux facteurs importants amplifient cette sous-optimalité. Le premier, nous le verrons au chapitre C, est la non prise en compte des contraintes techniques dans le calcul de la marge. Le deuxième est la sur-couverture du processus. Idéalement, le critère de risque 1% devrait être défini pour une marge de sécurité nulle. Ce qui n'est évidemment pas le cas en pratique puisque la marge de sécurité est en général de l'ordre de 1000 à 3000MW selon la saison. Enfin, et nous l'avons également déjà vu, nous accroissons ce phénomène de sur-protection en ne considérant que la marge à la pointe et en réduisant les opportunités d'arbitrage. Il est difficile de quantifier cette sous-optimalité mais, même faible, elle justifie pleinement que la résolution du problème global (5.3) soit un axe de recherche important à la suite de cette thèse ;
- l'**utilisation du parc de production comme levier de couverture** : Comme l'optimisation de la production ne fait plus partie du modèle d'optimisation, nous avons perdu la possibilité d'utiliser le parc de production pour satisfaire le critère de risque. Pourtant, il peut être bien plus rentable d'utiliser ses centrales que d'aller se fournir sur les marchés. Pour remédier à ce problème et sous la condition que nous soyons capables d'intégrer le calcul des capacités maximales de production thermique et hydraulique dans le problème (5.4) (et, en particulier, dans la contrainte en probabilité!), nous pouvons introduire deux commandes de sécurité supplémentaires considérées par les exploitants :
 - les déplacements d'arrêts en maintenance des centrales thermiques ;
 - l'utilisation du parc hydraulique.

Nous discuterons de cette éventualité au chapitre 9. Nous verrons notamment que, pour le cadre hydraulique, cette éventualité paraît pour l'instant très difficile à traiter. En revanche, dans le cadre de la capacité thermique, de premiers modèles ont été effectués et confortent l'importance de ces approches.

Pour conclure cette section, nous pouvons faire une analyse des propriétés mathématiques du problème (5.4). Nous rappelons que la marge est un processus stochastique connu et disponible sous forme de scénarios. La fonction objectif du problème est linéaire. La contrainte à l'intérieur de la contrainte en probabilité est linéaire ainsi que les diverses contraintes techniques sur les contrats. La taille est fonction du nombre de pas de temps (donc faible puisque égal à 12) et du nombre de contrats et de produits financiers disponibles. La taille devrait donc rester raisonnable. Cependant, ce problème comporte deux difficultés :

- la **contrainte en probabilité** : nous en avons déjà parlé au chapitre 3. La difficulté ma-

jeune de ce type de contraintes est qu'elles définissent un domaine réalisable généralement non convexe voire non connexe. Sous certaines hypothèses déjà discutées (séparabilité, log-concavité, . . .), nous pouvons cependant les transformer en contraintes convexes. D'autre part, il existe des approches pour approximer de telles contraintes. Nous rediscuterons par la suite des approches étudiées au cas par cas en fonction des hypothèses prises en compte ;

- des **variables entières** : le déclenchement ou l'achat/vente de contrats nécessite l'intégration de variables binaires. Mais l'introduction de variables entières dans une contrainte en probabilité, dans le cas général, rend cette éventualité très difficile.

La structure d'information joue également un rôle sur la nature du problème d'optimisation. En effet, si les variables sont en boucle ouverte (section 5.4), le problème (5.4) peut se ramener à un problème linéaire en nombres entiers. Si les variables sont en boucle fermée (chapitre 6), le problème devient si difficile qu'il n'existe pas à notre connaissance de résultats dans la littérature permettant de résoudre ce type de problèmes.

5.4 Modélisation du problème en boucle ouverte

Dans cette section, nous introduisons les notions de boucles ouverte et fermée et, ensuite, nous formulons le problème de couverture du risque physique dans le premier cas de la boucle ouverte.

5.4.1 Boucle ouverte et boucle fermée

Nous introduisons dans cette sous-sections deux façons différentes d'appréhender le problème, notamment au niveau de la séquence des décisions et des informations disponibles. Il s'agit en réalité d'une autre formulation de "problèmes à 2 étapes" et "problèmes multi-étapes" introduits dans le chapitre 2.

Dans ce chapitre, nous formulerons le problème à l'aide d'un schéma en boucle ouverte. Nous verrons au chapitre suivant une approche pour le cas de la boucle fermée.

Boucle ouverte

La notion de "boucle ouverte" est issue de l'automatique. Par définition, un contrôle en **boucle ouverte** est un contrôle optimal qui ne dépend pas a priori de l'état du système à un moment quelconque. Dans notre contexte, nous connaissons en t_0 une prévision de l'aléa pour chaque instant. Une optimisation en boucle ouverte consiste juste à prendre des décisions en t_0 pour tous les instants de la période d'étude de sorte à être optimal pour cette prévision.

Nous pouvons voir deux inconvénients majeurs à cette modélisation :

- l’anticipation des décisions : nous prenons des décisions avec une connaissance parfaite d’un futur qui n’a quasi aucune chance de se réaliser. Une telle approche est peu réaliste et très optimiste ;
- l’impossibilité de corriger : comme les décisions sont prises une fois pour toutes en t_0 , s’il apparaît des perturbations sur la prévision de l’aléa, il n’y a aucun moyen de corriger les commandes.

Cette modélisation est donc peu satisfaisante a priori. Mais, en pratique, un outil de couverture, à l’image d’un outil de gestion de production, a un grand intérêt à être lancé quotidiennement avec une mise à jour des prévisions. Dans ce cadre, l’utilisation d’une boucle ouverte n’est plus aussi problématique. En rebouclant quotidiennement, nous corrigeons les commandes prises en l’instant précédent en fonction d’une nouvelle prévision. Et l’anticipation est essentiellement problématique pour les instants les plus lointains. Si seules les commandes des premiers instants sont conservées, l’effet est limité. L’approche basée sur cette réoptimisation quotidienne peut donc s’avérer comme une approche satisfaisante en pratique et non dénuée de sens.

Boucle fermée

La **boucle fermée** s’oppose à la boucle ouverte, dans le sens qu’au lieu de connaître dès l’instant initial une prévision de l’aléa, nous ne connaissons la réalisation de l’aléa qu’à chaque instant. Les décisions deviennent séquentielles puisque la décision x_{t+1} ne sera prise qu’une fois l’aléa ξ_t connu. Les défauts de la boucle ouverte sont ainsi gommés : x_{t+1} peut être vue comme une correction de x_t par rapport à une réalisation ξ_t différente de la prévision espérée. D’autre part, le problème de l’anticipation est aussi traité puisque le futur est désormais inconnu. Le contrôle, aussi appelé feedback dans ce contexte, peut ainsi être vu comme une commande en temps réel. En optimisation, cette condition de boucle fermée se traduit par $x_{t+1} = f(\xi_1, \dots, \xi_t)$, c’est à dire que cette commande en t est une fonction de tout l’historique des réalisations de l’aléa. La commande en boucle fermée n’est donc plus une commande mais plutôt une loi de commande. Même si elle est plus réaliste qu’une boucle ouverte, elle a cependant le fort désavantage de compliquer notablement la nature des problèmes d’optimisation.

5.4.2 Formulation du problème avec les actifs financiers

Notations

L’introduction des divers contrats financiers nécessite un ensemble de notations :

Nom	Signification
J	Horizon de la période d'étude
Ω_J	Période d'étude : $\Omega_J = \{1, \dots, J\}$
Ω_J^-	Ensemble des jours ouvrables de Ω_J sans ceux de la dernière semaine calendaire
Ω_J^{We}	Ensemble des jours ouvrables de Ω_J d'une semaine dont le week-end appartient à la période d'étude
S_j	Ensemble des jours ouvrables de la semaine contenant le jour j
S_j^-	Ensemble des jours ouvrables de la semaine précédente à celle contenant le jour j

Les spécificités des contrats

Soit \mathcal{C} l'ensemble des contrats financiers considérés. Dans notre étude, \mathcal{C} est :

$$\mathcal{C} = \{\text{Spot Base, Spot Pointe, Week Ahead Base, Week Ahead Pointe, Week-End Ahead Base, Week-End Ahead Pointe}\}.$$

D'autre part, nous pourrions être emmenés à utiliser les sous-ensembles :

- $\mathcal{C}^S = \{\text{Spot Base, Spot Pointe}\}$;
- $\mathcal{C}^W = \{\text{Week Ahead Base, Week Ahead Pointe}\}$;
- $\mathcal{C}^{We} = \{\text{Week-End Ahead Base, Week-End Ahead Pointe}\}.$

Nous définissons $q_j^{A,c}$ (resp. $q_j^{V,c}$) la quantité d'électricité achetée (resp. vendue) à la date j à l'aide du contrat $c \in \mathcal{C}$. Pour chaque contrat c , nous spécifions son ensemble de faisabilité U_c . Nous pouvons détailler ces ensembles par type de contrat :

- Pour le Spot (Base ou Pointe) :

$$U_S = \left\{ q_j^{A,S}, q_j^{V,S} : 0 \leq q_j^{A,S} \leq \bar{q}_j^{A,S}, 0 \leq q_j^{V,S} \leq \bar{q}_j^{V,S}, \forall j = 1, \dots, J \right. \\ \left. q_j^{A,S} = 0, q_j^{V,S} = 0 \text{ si } j \in \{\text{Samedi, Dimanche}\} \right\}$$

Nous limitons uniquement les quantités d'achats/ventes sur le marché Spot pour deux raisons. D'une part, la profondeur des marchés est relativement faible. D'autre part, EDF est un "price maker". Cela signifie qu'EDF est un acteur si important sur le marché que ses actions influent sur les prix (à la hausse ou à la baisse). Les quantités négociées sont donc limitées de manière à pouvoir négliger cette influence sur les prix.

- Pour les Week-Ahead (W) (Base ou Pointe) :

$$U_W = \left\{ q_j^{A,W}, q_j^{V,W} : 0 \leq q_j^{A,W} \leq \bar{q}_j^{A,W}, 0 \leq q_j^{V,W} \leq \bar{q}_j^{V,W}, \forall j = 1, \dots, J \right. \\ \left. q_j^{A,W} = 0, q_j^{V,W} = 0 \text{ si } j \in \{\text{Samedi, Dimanche}\} \right\}$$

Les variables d'achats/ventes en Week Ahead ne possèdent qu'un unique indice temporel. Il faut tout de même garder en tête qu'il existe une contrainte sous-jacente imposant que, par

exemple, la quantité $q_j^{A,WA}$ doit être livrée en quantité égale à chaque jour ouvrable de la semaine suivante.

– Pour les Week End-Ahead (We) (Base ou Pointe) :

$$U_{We} = \left\{ q_j^{A,We}, q_j^{V,We} : 0 \leq q_j^{A,We} \leq \bar{q}_j^{A,We}, 0 \leq q_j^{V,We} \leq \bar{q}_j^{V,We}, \forall j = 1, \dots, J \right. \\ \left. q_j^{A,We} = 0, q_j^{V,We} = 0 \text{ si } j \in \{\text{Samedi, Dimanche}\} \right\}$$

A l'image des Week Ahead, les quantités reçues les samedis et dimanches doivent être identiques.

La gestion des Options Tarifaires

D'autres outils de couverture sont les Options Tarifaires (OT). Les puissances et les coûts sont contractualisés. Les décisions correspondantes à l'utilisation des options tarifaires sont donc binaires : soit nous déclenchons ces options et nous effaçons la quantité $q_{i,j}^{OT}$, soit nous ne les déclenchons pas. Soit la variable de décision binaire e_j^k définie par :

$\forall j = 1, \dots, J-1, \forall k = 1, \dots, |OT| :$

$$e_j^k = \begin{cases} 1 & \text{si l'OT } k \text{ est déclenchée en } j \text{ pour effacement en } j+1 \\ 0 & \text{sinon.} \end{cases}$$

Pour la gestion des OT, nous devons également introduire des contraintes particulières. La première contrainte impose qu'une option tarifaire ne puisse être déclenchée qu'une fois :

$$\forall k = 1, \dots, |OT|, \sum_{j=1}^{J-1} e_j^k = 1.$$

D'autre part, en gestion hebdomadaire, les exploitants disposent en général d'un nombre maximal N_{OT} d'options tarifaires déclenchables sur la période d'étude. Ce nombre est fixé par les équipes annuelles. Cette contrainte s'écrit :

$$\sum_{k=1}^{|OT|} \sum_{j=1}^{J-1} e_j^k \leq N_{OT}.$$

La fonction économique

L'objectif du problème est de minimiser le coût total de la couverture, soit l'écart minimal entre les dépenses et les revenus. Nous introduisons les notations suivantes : soit $p_j^{A,c}$ (resp. $p_j^{V,c}$) le

prix d'achat (resp. de vente) du contrat $c \in \mathcal{C}$ à la date j . Ces prix étant aléatoires, nous pouvons discrétiser cet aléa sous la forme de scénarios $p_{j,n}^{A,c}$ ou $p_{j,n}^{V,c}$. Enfin, les prix des options tarifaires sont fixés et contractualisés. Nous noterons $p_{i,j}^{OT}$ le prix de déclenchement de l'option tarifaire i à la date j pour effacement à $j+1$. A l'aide des notations de la section 5.4.2, la fonction objectif s'écrit :

$$C(q_j^{A,c}, q_j^{V,c}, e_j^k) = \mathbb{E} \left[\sum_{j=1}^{J-1} \sum_{c \in \mathcal{C}^S} (p_j^{A,c} q_j^{A,c} - p_j^{V,c} q_j^{V,c}) + 5 \sum_{j \in \Omega_J^-} \sum_{c \in \mathcal{C}^W} (p_j^{A,c} q_j^{A,c} - p_j^{V,c} q_j^{V,c}) \right. \\ \left. + 2 \sum_{j \in \Omega_J^{We}} \sum_{c \in \mathcal{C}^{We}} (p_j^{A,c} q_j^{A,c} - p_j^{V,c} q_j^{V,c}) + \sum_{j=1}^{J-1} \sum_{k=1}^{|OT|} p_{k,j}^{OT} q_{k,j}^{OT} e_j^k \right].$$

Les coefficients 5 et 2 étant le nombre de jours de livraison de chaque contrat.

Tous les scénarios de prix étant supposés équiprobables, nous pouvons récrire l'objectif :

$$C(q_j^{A,c}, q_j^{V,c}, e_j^k) = \frac{1}{N} \sum_{n=1}^N \left[\sum_{j=1}^{J-1} \sum_{c \in \mathcal{C}^S} (p_{j,n}^{A,c} q_j^{A,c} - p_{j,n}^{V,c} q_j^{V,c}) + 5 \sum_{j \in \Omega_J^-} \sum_{c \in \mathcal{C}^W} (p_{j,n}^{A,c} q_j^{A,c} - p_{j,n}^{V,c} q_j^{V,c}) \right. \\ \left. + 2 \sum_{j \in \Omega_J^{We}} \sum_{c \in \mathcal{C}^{We}} (p_{j,n}^{A,c} q_j^{A,c} - p_{j,n}^{V,c} q_j^{V,c}) + \sum_{j=1}^{J-1} \sum_{k=1}^{|OT|} p_{k,j}^{OT} q_{k,j}^{OT} e_j^k \right].$$

Nous pouvons simplifier cette écriture en introduisant le prix moyen sur l'ensemble des scénarios au pas de temps j :

$$\bar{p}_j = \frac{\sum_{n=1}^N p_{j,n}}{N}.$$

Le coût devient enfin :

$$C(q_j^{A,c}, q_j^{V,c}, e_j^k) = \sum_{j=1}^{J-1} \sum_{c \in \mathcal{C}^S} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + 5 \sum_{j \in \Omega_J^-} \sum_{c \in \mathcal{C}^W} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) \\ + 2 \sum_{j \in \Omega_J^{We}} \sum_{c \in \mathcal{C}^{We}} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + \sum_{j=1}^{J-1} \sum_{k=1}^{|OT|} p_{k,j}^{OT} q_{k,j}^{OT} e_j^k.$$

La contrainte de gestion du risque physique

La contrainte de gestion du risque est une contrainte en probabilité. Elle impose un bilan journalier de puissance disponible, qu'elle soit produite par le parc ou achetée sur les marchés, supérieur

à un seuil de sécurité sauf éventuellement sur les $p\%$ de scénarios défavorables. La contrainte s'écrit :

$$\begin{aligned} \forall j \in \Omega_J : \text{Prob} \left(\tilde{M}_j + \sum_{c \in \mathcal{C}^S} (q_{j-1}^{A,c} - q_{j-1}^{V,c}) + \sum_{k \in S_j^-} \sum_{c \in \mathcal{C}^W} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \notin \{\text{Sam}, \text{Dim}\}\}} \right. \\ \left. + \sum_{k \in S_j} \sum_{c \in \mathcal{C}^{We}} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \in \{\text{Sam}, \text{Dim}\}\}} + \sum_{k=1}^{|OT|} q_{k,j-1}^{OT} e_{j-1}^k \geq M_S \right) \geq 1 - p \end{aligned} \quad (5.5)$$

Les variables du problème sont en boucle ouverte donc sans aucune dépendance par rapport à un aléa. La seule variable aléatoire de la contrainte (5.5) est donc la marge de production. Nous nous situons dans le cas séparable où l'aléa n'est pas en produit avec une variable de décision et est scalaire.

Prenons le cas d'une contrainte générale de la forme $\text{Prob}(M + X \geq M_S) \geq 1 - p$, soit encore $\text{Prob}(-M \leq X - M_S) \geq 1 - p$, où M est une variable aléatoire scalaire et $X, M_S \in \mathbb{R}$.

Notons :

- $F_M(z) = \text{Prob}(M \leq z)$ la **fonction de répartition** de la variable aléatoire M ;
- $q_M(p) = \inf\{\tau | F_M(\tau) \geq p\}$ le **quantile** de la variable aléatoire M à l'ordre p .

Si M est une variable aléatoire scalaire, alors nous avons l'équivalence suivante :

$$F_M(z) \geq p \Leftrightarrow z \geq q_M(p).$$

La contrainte en probabilité est donc équivalente à la contrainte linéaire :

$$X - M_S \geq q_{-M}(1 - p).$$

En appliquant cette transformation à notre contrainte de marge, nous transformons une contrainte en probabilité en une contrainte linéaire standard :

$$\begin{aligned} \forall j \in \Omega_J : q_{\tilde{M}}^j(p) + \sum_{c \in \mathcal{C}^S} (q_{j-1}^{A,c} - q_{j-1}^{V,c}) + \sum_{k \in S_j^-} \sum_{c \in \mathcal{C}^W} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \notin \{\text{Sam}, \text{Dim}\}\}} \\ + \sum_{k \in S_j} \sum_{c \in \mathcal{C}^{We}} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \in \{\text{Sam}, \text{Dim}\}\}} + \sum_{k=1}^{|OT|} q_{k,j-1}^{OT} e_{j-1}^k \geq M_S, \end{aligned} \quad (5.6)$$

où $q_{\tilde{M}}^j(p)$ est le quantile à $1 - p$ de la variable aléatoire \tilde{M} au pas de temps j .

Il existe d'autres approches pouvant approximer des contraintes en probabilité par des contraintes convexes pour des problèmes en boucle ouverte. Nous pouvons notamment citer les approximations par scénarios ([4],[8]), les approximations à l'aide d'approches probabilistes [9] et évidemment l'optimisation robuste (cf. 3). A noter qu'il existe aussi des premières approches pour résoudre des problèmes avec variables binaires dans la contrainte en probabilité [7].

Formulation générale

En récapitulant les modélisations précédentes, nous pouvons écrire le problème général :

$$\begin{aligned}
 \min_{q^A, q^V, e} \quad & \sum_{j=1}^{J-1} \sum_{c \in \mathcal{C}^S} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + 5 \sum_{j \in \Omega_J^-} \sum_{c \in \mathcal{C}^W} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) \\
 & + 2 \sum_{j \in \Omega_J^{We}} \sum_{c \in \mathcal{C}^{We}} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + \sum_{j=1}^{J-1} \sum_{k=1}^{|OT|} p_{k,j}^{OT} q_{k,j}^{OT} e_j^k. \\
 \text{s.t.} \quad & \left\{ \begin{aligned} & q_M^j(p) + \sum_{k \in S_j^-} \sum_{c \in \mathcal{C}^W} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \notin \{\text{Sa}, \text{Di}\}\}} + \sum_{k \in S_j} \sum_{c \in \mathcal{C}^{We}} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \in \{\text{Sa}, \text{Di}\}\}} \\ & + \sum_{c \in \mathcal{C}^S} (q_{j-1}^{A,c} - q_{j-1}^{V,c}) + \sum_{k=1}^{|OT|} q_{k,j-1}^{OT} e_{j-1}^k \geq M_S, \quad \forall j \in \Omega_J \\ & \sum_{j=1}^{J-1} e_j^k = 1, \quad \forall k = 1, \dots, |OT| \\ & \sum_{k=1}^{|OT|} \sum_{j=1}^{J-1} e_j^k \leq N_{OT} \\ & \forall c \in \mathcal{C}^S, \quad q^{A,c}, q^{V,c} \in U_S \\ & \forall c \in \mathcal{C}^W, \quad q^{A,c}, q^{V,c} \in U_W \\ & \forall c \in \mathcal{C}^{We}, \quad q^{A,c}, q^{V,c} \in U_{We} \end{aligned} \right.
 \end{aligned} \tag{5.7}$$

Ce problème est un problème linéaire à variables mixtes. La taille du problème demeure raisonnable puisqu'elle dépend du nombre de pas de temps (a priori 12), du nombre de produits financiers et du nombre de contrats d'effacement. Nous avons :

- Nombre de variables : $2(J-1)|\mathcal{C}^S| + 2|\Omega_J^-||\mathcal{C}^W| + 2|\Omega_J^{We}||\mathcal{C}^{We}| + (J-1)|OT|$;
- Nombre de contraintes (sans les contraintes de bornes) : $|\Omega_J| + |OT| + 1$.

5.4.3 Calcul des quantiles

Pour calculer les quantiles, nous procédons en deux étapes : nous estimons la densité de probabilité du processus de marge à partir des scénarios de marge disponibles puis nous calculons le quantile souhaité de cette distribution.

Nous étudions et illustrons le calcul de densité au chapitre C. En pratique, les densités sont estimées via la fonction `ksdensity` du logiciel Matlab, utilisant la méthode des noyaux.

Les quantiles sont calculés à l'aide de la fonction Matlab `quantile`. Elle crée implicitement un histogramme de la densité et elle calcule le quantile demandé par interpolation linéaire.

5.4.4 Résultats

Nous appliquons l'approche exposée sur la situation décrite sur la figure 5.2. A partir de 75 000 scénarios de marge, nous estimons une chronique de quantiles à 1% sur une période de 12 jours. Ces données initiales montrent, qu'entre les deux dimanches, le critère de risque n'est pas satisfait. Pour remédier à cela, nous considérons le marché Spot (1000 MW à l'achat et 500 MW à la vente autorisés) et un contrat Week Ahead (400 MW autorisés à l'achat et à la vente), disponibles tous deux en Base et en Pointe. Nous considérons de plus un contrat Week-End Ahead Base (200 MW disponibles à l'achat et à la vente) et 10 options tarifaires (de puissance effaçable de 1000, 1500, 2000 ou 3000 MW). La marge de sécurité est fixée à 3000 MW pour chaque jour.

Ce problème de taille raisonnable (570 contraintes, 620 variables) est résolu en moins d'une seconde sur un PC 1.6 GHz à l'aide du kit de modélisation et de résolution GLPK. Les commandes de couverture obtenues (Fig.5.3) permettent de satisfaire le critère (Fig.5.2).

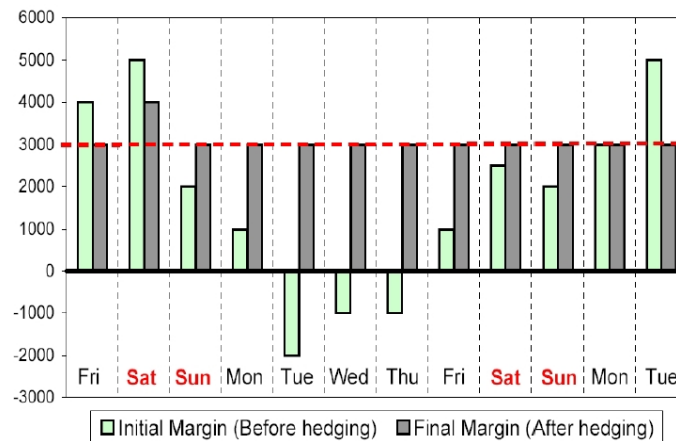


FIG. 5.2 – Marge de production avant et après la couverture

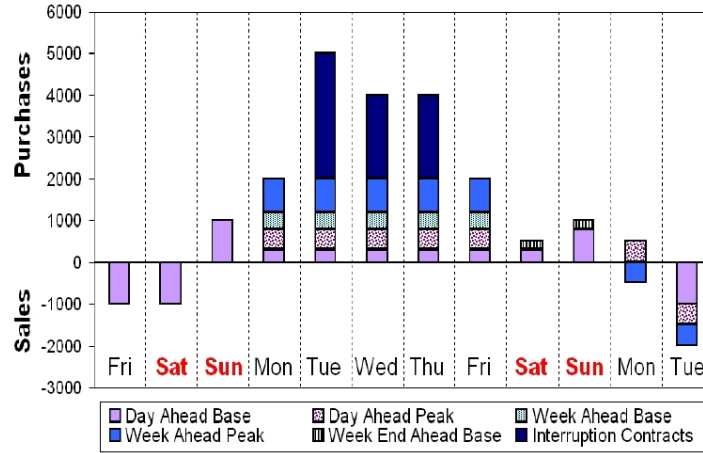


FIG. 5.3 – Commandes de couverture issues de l'optimisation

5.5 Conclusion

Dans cette section, nous avons étudié le problème de couverture contre le risque physique à l'horizon hebdomadaire. Après une analyse des outils existants et des possibilités offertes, nous nous sommes dirigés vers des approches simplifiées, basées sur la marge de production et proches des pratiques appliquées en exploitation.

Nous avons formulé le problème sous la forme d'un problème linéaire en nombres entiers avec des variables en boucle ouverte. Cette approche permet de donner des réponses à l'exploitant mais nous pouvons chercher à améliorer le processus. Une première façon de faire est de rendre le processus de décision glissant et de réitérer l'optimisation tous les jours à la vue de nouvelles prévisions des aléas. Une autre façon de faire est de rendre les décisions plus réalistes en les rendant adaptatives, mais toutefois non-anticipatives. Ce sont des variables en boucle fermée. Nous étudierons donc au chapitre suivant des stratégies d'optimisation prenant en compte l'information disponible à chaque instant.

Bibliographie

- [1] L. Andrieu, R. Apparigliato, A. Lissier, and A. Tang. Stochastic Optimization under Risk Constraints : Application to Hedging Problem in Electrical Industry. *11th International Conference on Stochastic Programming, Vienna*, August 27-31 2007.
- [2] L. Andrieu, R. Apparigliato, A. Lissier, and A. Tang. Stochastic Optimization under Risk Constraints : Application to Hedging Problem in Electrical Industry. *Power Systems Management Conference, Athènes*, June 5-8 2007.
- [3] R. Apparigliato and R. Zorgati. Couverture et Arbitrage sur les Marchés en Gestion Hebdomadaire de la Production Electrique. Partie II : Formulation du Problème de Couverture. Technical report, EDF R&D, Dept OSIRIS, A paraître.

- [4] P. Bonami and M. A. Lejeune. An Exact Solution Approach for Portfolio Optimization Problems under Stochastic and Integer Constraints. Technical report, 2007.
- [5] R. Henrion. Introduction to Chance-Constrained Programming. *Tutorial paper for the Stochastic Programming Community Home Page*, 2004.
- [6] R. Henrion. Optimization Problems with Probabilistic Constraints. *10th International Conference on Stochastic Programming, Tucson*, 2004.
- [7] Olivier Klopfenstein. Tractable algorithms for chance-constrained combinatorial problems. Technical report, France Telecom R&D, 38-40 rue du Général Leclerc, 92130 Issy-les-Moulineaux, France, February 15, 2007.
- [8] A. Nemirovski and A. Shapiro. Scenario Approximations of Chance Constraints. Preprint, School of Industrial and Systems Engineering, Georgia Institute of Technology, 2004.
- [9] A. Nemirovski and A. Shapiro. Convex Approximations of Chance Constrained Programs. *SIAM Journal on Optimization*, 17 :969–996, 2006.
- [10] A. Prékopa. On Probabilistic Constrained Programming. In Princeton University Press, editor, *Proceedings of the Princeton Symposium on Mathematical Programming*, pages 113–138, 1970.
- [11] A. Prékopa. *Stochastic Programming*. Kluwer, Dordrecht, Boston, 1995.
- [12] A. Prékopa. Probabilistic Programming. In A. Ruszczyński and A. Shapiro, editors, *Handbooks in Operations Research and Management Science*, volume 10. North-Holland, 2003.
- [13] F. Romanet. Méthodologie Générale de Résolution Employée dans l’Optimiseur SOPRANO. Technical report, EDF R&D, Dept MOSSE, MM-31/02/018/A, 2002.
- [14] F. Romanet. SOPRANO : Présentation et Spécifications Fonctionnelles. Technical report, EDF R&D, Dept MOSSE, MM-31/02/017/A, 2002.
- [15] G. Vignal. Spécifications Fonctionnelles du Placement des Options Tarifaires dans le Code APOGEE 12 jours. Technical report, EDF R&D, Dept OSIRIS, MM-37/03/010/A, 2003.
- [16] G. Vignal. Modèle APOGEE. Vers une Nouvelle Formulation des Sous-Problèmes Hydrauliques. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/003/A, 2004.
- [17] G. Vignal. Spécifications Générales du Modèle APOGEE 12 jours - Description des Données et des Résultats. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/010/A, 2004.
- [18] R. Zorgati and P. Thomas. Formation au Modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d’une Vallée sur Critère Economique. I. Formulation. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/015/A, 2004.
- [19] R. Zorgati and P. Thomas. Formation au Modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d’une Vallée sur Critère Economique. II. Résolution. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/016/A, 2004.
- [20] R. Zorgati and P. Thomas. Les Marchés dans OPHELIE : Synthèse des Modélisations. Technical report, EDF R&D, Dept OSIRIS, HR-37/03/017/A, 2004.
- [21] R. Zorgati, W. Van Ackooij, J.-M. Luel, P. Thomas, M. Uchanski, and K. Shea. LEA : A Cluster Intensive Simulation Software For Unit Commitment. *Proceedings of the 7th LCI High Performance Computing Conference, Norman, Oklahoma, (USA)*, May 02-04 2006.

Chapitre 6

Boucle fermée : la programmation stochastique avec Règles de Décision constantes par morceaux

Sommaire

6.1	Introduction	174
6.2	Programmation stochastique avec règles de décision constantes par morceaux	176
6.2.1	Formalisme	176
6.2.2	De la programmation stochastique à la SPSDR...	177
6.2.3	L'approche SPSDR en détails	181
6.2.4	Résultats obtenus sur le problème de chaînes d'approvisionnement	186
6.3	Adaptations de l'approche pour le problème de couverture	186
6.3.1	Aléa multidimensionnel	186
6.3.2	Contraintes sur l'état	188
6.3.3	Comparaisons numériques entre les heuristiques de calcul du problème p -médiannes	188
6.4	Formulation simplifiée du problème	190
6.5	Ecriture du problème de couverture avec SPSDR	192
6.6	Résultats numériques : pré-requis	193
6.6.1	Jeu d'essais	193
6.6.2	Modèles simples de comparaison	194
6.7	Résultats numériques avec SPSDR	197
6.7.1	Effets du nombre d'ensembles de décisions et de la taille des experts	198
6.7.2	Effets du nombre d'experts	200
6.7.3	Introduction d'un coefficient d'immunité	201
6.7.4	Synthèse des résultats	203
	Bibliographie	205

Dans ce chapitre, nous proposons d'appliquer au problème de couverture présenté au chapitre précédent une approche permettant d'aborder des problèmes stochastiques à plusieurs étapes : la Programmation Stochastique avec règles de décision constantes par morceaux. Cette approche n'a été introduite à l'heure actuelle que dans le cadre des chaînes d'approvisionnement [14]. Pour l'appliquer à notre problème de couverture, nous avons dû notamment adapter l'approche pour gérer des contraintes sur les états, en particulier en intégrant des coefficients de sécurité, et surtout en étendant l'approche au cas d'aléas multidimensionnels. Les résultats montrent, qu'en jouant sur les paramètres sur les modèles, il est possible de contrôler les violations de contraintes et d'obtenir de très bonnes performances sur les coûts. Les travaux lors de cette thèse ont donné lieu à une présentation au congrès Stochastic Programming [1].

6.1 Introduction

L'approche étudiée a été proposée par Julien Thénier et Jean-Philippe Vial dans l'article "Stochastic Programming with Step Decision Rules : a heuristic approach" [14]. Cette étude a d'ailleurs été effectuée en collaboration avec les auteurs. La présentation qui suit de cette heuristique est également fondée sur [14]. Enfin, pour faire référence à cette approche, nous conserverons tout au long de ce chapitre l'abréviation SPSDR issue du nom anglais précédemment introduit.

Pour intégrer la notion d'incertitude au sein de modèles d'optimisation, de nouvelles approches ne cessent d'apparaître dans la littérature. Cependant, elles sont généralement appliquées dans le cadre statique. Or, les problèmes sont généralement "dynamiques", dans le sens où une décision doit être prise à chaque pas de temps sur la base de l'état réel du système (en particulier des réalisations des aléas). Malheureusement, ce type de problèmes sont difficiles. Nous avons vu au chapitre 2 les deux principales difficultés : une complexité rendant ces problèmes non traitables numériquement et des calculs d'espérance qui croissent exponentiellement avec la dimension du processus et du nombre de pas de temps. Il est donc raisonnable de privilégier des approches certes sous-optimales mais donnant des résultats suffisamment performants en des temps de calculs compatibles avec une application industrielle.

L'heuristique proposée dans ce chapitre, la SPSDR, est dans cet esprit. Elle permet de déterminer des politiques pour des problèmes de contrôle stochastique sous contraintes dynamiques linéaires et variables d'état non contraintes. Nous supposons également que les décisions prises n'influent pas sur le processus stochastique. Cette heuristique est basée sur le concept de règle de décision qui n'est rien d'autre qu'un mécanisme qui sélectionne la décision à prendre en un instant sur la base de l'information disponible. L'ensemble des réalisations possibles étant vaste (voire infini), nous considérons seulement une sous-classe de règles de décision, les *règles de décision constantes par morceaux*, qui ne sont que des fonctions en escalier sur l'espace des réalisations du processus stochastique. Une telle règle de décision a donc deux composantes : le support de la fonction en escalier qui est une partition de l'ensemble des réalisations du processus jusqu'à un instant t et la valeur des variables de décision en t sur chaque élément de la partition. Un élément de la partition est appelé un *ensemble de décisions*. Quand les ensembles de décisions sont fixés, le problème de trouver une règle de décision optimale consiste à assigner des décisions à chaque

ensemble de décisions.

La programmation stochastique multi-étapes est connue pour souffrir d’une augmentation exponentielle du nombre de variables (de décisions et d’états) lorsque le nombre d’étapes augmente linéairement. Ce résultat est dû à la contrainte de raffinement qui contraint l’arbre à avoir de plus en plus de branches au cours du temps, modélisant ainsi une incertitude croissante avec le temps. Pour contourner ce problème, la programmation stochastique avec règles de décision constantes par morceaux n’est plus contrainte par ce raffinement. Cette hypothèse peut introduire une inconsistance dans les décisions qui peut être interprétée comme une perte de mémoire du décideur. La décision ainsi obtenue est sous-optimale mais les temps de calculs (et les éventuels problèmes de mémoire informatique) sont potentiellement réduits. D’autre part, le processus reste clairement non-anticipatif puisque la partition des ensembles d’incertitude n’est créée que sur la base d’informations passées.

Une instance classique de la SPSDR peut donc être résumée comme suit. Nous générons un premier ensemble de scénarios modélisant le processus stochastique sur lequel nous construisons, pour chaque pas de temps, une partition en ensembles de décisions basée sur l’information passée. Cette construction est obtenue par résolution d’un problème de p -médianes où la distance choisie est la fonction de Kantorovitch (comme dans [10, 7]). La valeur assignée sur chaque ensemble de décisions est finalement obtenue en optimisant l’espérance du coût sur l’ensemble de scénarios. Il est possible d’améliorer la technique SPSDR en considérant une combinaison convexe de règles de décision indépendantes. Cela correspond à refaire l’implémentation précédente sur plusieurs tirages de scénarios indépendants. Nous pouvons interpréter cette combinaison comme l’avis de plusieurs experts indépendants face au problème à traiter. Cette idée d’experts est directement inspirée de [9].

Comme nous l’avons déjà évoqué, nous ne pouvons résoudre des problèmes à plusieurs étapes avec une précision garantie. Cependant, nous pouvons estimer la qualité de la solution dans une phase de validation. Pour cela, nous étendons la solution obtenue sur un pool de scénarios plus grand que celui utilisé pour l’optimisation des règles. Nous pouvons ensuite calculer des indicateurs comme les coûts moyens, les écarts-types sur les coûts, les coûts minimal et maximal,...

Le problème que nous souhaitons traiter est le problème de couverture dans la gestion des risques physiques présenté au chapitre 5. Contrairement aux problèmes traités dans [14], notre cas d’étude prend en compte un aléa multidimensionnel (marge, prix spot, prix à terme). Nous avons donc ajusté l’approche à nos besoins (voir section 6.3). Pour cela, nous créons une partition pour chacune des dimensions de l’aléa et la règle finale est une combinaison des règles issues de l’optimisation de chacune des dimensions. La règle finale a donc comme support le produit cartésien des partitions de chaque dimension de l’aléa.

Il n’a pas été possible dans le temps imparti de comparer les performances numériques de SPSDR à d’autres méthodes d’optimisation dans l’incertain sur notre problème. Nous nous référons toutefois à [14] où les résultats obtenus sur un problème de chaîne d’approvisionnement sont tout à fait encourageants par rapport à la programmation stochastique classique sur arbre (PSA) et à l’optimisation robuste. Cet article souligne que, pour ce problème et le jeu de données choisi, la PSA commet une erreur importante de prédiction (de l’ordre de 12%), ce qui représente l’écart entre le coût optimal annoncé après optimisation sur l’arbre et le coût optimal après validation sur un autre échantillon de scénarios. En comparaison, l’erreur de prédiction de SPSDR est

inférieure à 1%. Enfin, le coût optimal est amélioré par SPSDR de presque 5% en simulation par rapport à celui obtenu avec la programmation stochastique. L'optimisation robuste fournit une stratégie de risque efficace mais très coûteuse. Son coût est donc une borne maximale.

Nous présenterons deux séries de résultats sur notre problème : une série pour illustrer le comportement de SPSDR en fonction des paramètres initiaux (nombre d'experts, taille des experts, nombre d'ensembles de décisions) et une série pour comparer les résultats avec les modèles en boucle ouverte du chapitre 5 et leur version robuste.

6.2 Programmation stochastique avec règles de décision constantes par morceaux

Dans cette section, nous présentons l'heuristique telle qu'elle a été introduite dans l'article de référence [14]. Les adaptations effectuées dans le cadre de la thèse seront présentées dans la section 6.3.

6.2.1 Formalisme

Plaçons nous d'abord dans le cadre général pour lequel cette approche a été élaborée. Nous supposons le problème déterministe général, d'horizon T , suivant :

$$\begin{aligned} \min_{u_1, \dots, u_{T-1}} \quad & \phi(x_1, \dots, x_T, u_1, \dots, u_{T-1}) \\ \text{s.t.} \quad & \begin{cases} x_t = A_t x_{t-1} + B_t u_{t-1} + C_t, \quad \forall t = 2, \dots, T \\ g_t(x_1, \dots, x_t, u_1, \dots, u_t) \leq 0, \quad \forall t = 1, \dots, T-1 \\ x_1 = \bar{x}_1, \end{cases} \end{aligned} \tag{6.1}$$

où u_t est le contrôle du problème, x_t la variable d'état, A_t, B_t et C_t des matrices de taille adéquate et \bar{x}_1 est une valeur initiale fixée. Nous supposons les contraintes g_t convexes conjointement en (x_1, \dots, x_t) et en (u_1, \dots, u_t) ; de même, la fonction objectif ϕ est convexe conjointement en (x_1, \dots, x_T) et (u_1, \dots, u_{T-1}) . Pour simplifier, nous considérerons une fonction-objectif séparable :

$$\phi(x_1, \dots, x_T, u_1, \dots, u_{T-1}) = \sum_{t=1}^{T-1} f_t(x_t, u_t) + f_T(x_T).$$

Le problème devient stochastique lorsque les composantes de (6.1) dépendent des réalisations $\{\xi_t\}_{t=1}^T$ d'un processus stochastique $\tilde{\xi}$. Nous ferons l'hypothèse suivante :

Hypothèse 6.1 *Les réalisations du processus stochastique sont indépendantes du processus de décision.*

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

Le processus stochastique est muni d'un espace probabilisé (Ξ, \mathcal{S}, μ) où Ξ est l'espace des réalisations de $\tilde{\xi}$, \mathcal{S} est une tribu de parties de Ξ et μ une probabilité sur \mathcal{S} . Nous pouvons introduire les espaces probabilisés $(\Xi_t, \mathcal{S}_t, \mu_t)_{t=1}^T$ des réalisations partielles $\sigma_t = (\xi_1, \dots, \xi_t)$. Les tribus \mathcal{S}_t engendrées par les variables aléatoires σ_t forment une filtration : nous avons $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \dots \subset \mathcal{S}_T$.

Pour résoudre le problème, nous faisons les hypothèses suivantes :

- nous nous plaçons dans une situation dite de “Hasard-Décision” où chaque décision n'est prise qu'après avoir pris connaissance de la réalisation de l'aléa ;
- l'espace Ξ_1 est un singleton pour que le problème ne soit pas décomposable en sous-problèmes indépendants ;
- la contrainte g_t n'est dépendante que des contrôles ;
- l'aléa est uni-dimensionnel.

D'autre part, une décision u_t est prise en toute connaissance de la filtration σ_t . u_t est donc un contrôle en boucle fermée $u_t(\sigma_t)$ puisque \mathcal{S}_t -mesurable. L'état x_t étant dépendant de ξ_t et de $u_{t-1}(\sigma_{t-1})$, il devient lui aussi contingent à σ_t . Enfin, nous pourrions borner les états x_t mais aucune garantie ne pourra être fournie sur le respect de ces bornes en simulation.

La version stochastique de ce problème devient donc :

$$\begin{aligned} \min_u \quad & \mathbb{E} \left[\sum_{t=1}^{T-1} f_t(x_t(\sigma_t), u_t(\sigma_t), \xi_t) + f_T(x_T(\sigma_T), \xi_T) \right] \\ \text{s.t.} \quad & \begin{cases} x_t(\sigma_t) = A_t(\xi_t)x_{t-1}(\sigma_{t-1}) + B_t(\xi_t)u_{t-1}(\sigma_{t-1}) + C_t(\xi_t), \\ g_t(u_t(\sigma_t)) \leq 0, \quad \forall t = 1, \dots, T-1 \\ x_1 = \bar{x}_1. \end{cases} \end{aligned} \tag{6.2}$$

6.2.2 De la programmation stochastique à la SPSDR...

Comme le processus stochastique est indépendant des décisions et que l'espace des réalisations est fini, nous pouvons écrire le problème (6.2) comme un problème de programmation stochastique et le résoudre comme un problème standard (de grande taille) de programmation mathématique.

La première étape est de faire des **analyses statistiques** sur le processus stochastique (qui est unidimensionnel) afin de l'identifier et d'estimer ses paramètres. Cette étape est importante puisque, même si le modèle d'optimisation est performant, une mauvaise estimation ou un mauvais calage peut enlever toute pertinence au modèle. Dans le cas où cette étape est bien effectuée, nous pouvons créer un générateur de scénarios échantillonnant le processus stochastique. Si le processus est continu, tous les scénarios doivent être différents à partir de l'étape 2 (la première étape est fixée, identique à tous les scénarios générés car Ξ_1 est un singleton). Concrètement, nous disposons d'un “peigne” de scénarios :

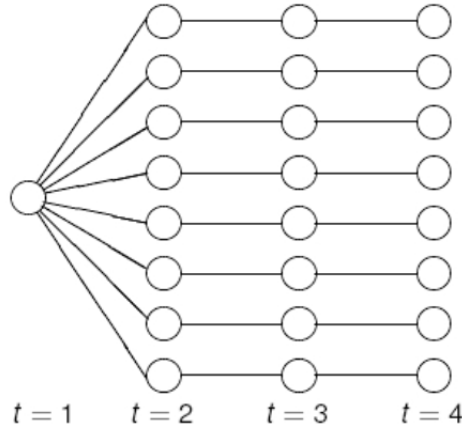


FIG. 6.1 – Représentation d'un peigne de scénarios

Le principal problème dans la prise de décisions multi-étapes dans l'incertain est de savoir comment l'incertitude est dévoilée. Si à un instant t nous apprenons tout le futur, le problème devient un problème à 2 étapes où nous prenons finalement un ensemble de décisions avant t et un ensemble de décisions pour tous les instants après t . C'est exactement la situation du peigne de scénarios de la figure 6.1 où, dès que la réalisation de l'aléa en $t = 2$ est connue, nous sommes capables d'anticiper tout le futur et donc de trouver une solution optimale pour tous les pas de temps suivants sur chaque scénario.

Ici, ce qui nous intéresse est non pas de résoudre un tel problème à 2 étapes mais plutôt de résoudre un problème à t étapes : à chaque t une réalisation de l'aléa est connue mais nous ne pouvons prendre une décision que pour l'instant t puisque la période $[t + 1, \dots, T]$ demeure toujours incertaine. La représentation du peigne n'est pas adéquate dans ce cas. Pour corriger cela, nous effectuons une **agrégation** de scénarios : cela consiste à imposer en chaque instant une même décision sur tous les scénarios qui sont "proches" (cette proximité est définie selon une certaine distance à choisir). Au lieu d'avoir ainsi une commande par instant et par scénario, nous avons désormais une commande par instant et par agrégat de scénarios. Cet agrégat est appelé **nœud**. Cependant, si à partir de chaque nœud, nous ne voulons pas être en totale anticipativité, il doit y avoir plusieurs futurs possibles : chaque nœud de t doit être ainsi relié à plusieurs nœuds de $t + 1$. C'est une **contrainte de raffinement** que l'agrégation doit prendre en compte. Nous venons ainsi de construire un **arbre de scénarios** :

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

Cependant, le principal défaut de la programmation stochastique est la taille du problème équivalent obtenu. Comme, en chaque instant, nous partitionnons les scénarios de sorte à être un raffinement de la partition de l’instant précédent, la taille de l’arbre croît exponentiellement quand le nombre de périodes augmente linéairement. Shapiro et Nemirovski [13] avancent que les “problèmes multipériodiques, même linéaires, avec recours complets sont numériquement difficiles à calculer même pour une précision moyenne de la solution”. Evidemment, cela n’empêche pas d’être performant sur certains problèmes mais cela justifie la recherche d’heuristiques donnant une bonne approximation pour des temps de calculs raisonnables. C’est dans cet esprit que nous proposons d’appliquer la **programmation stochastique avec règles de décision continues par morceaux** (Stochastic Programming with Step Decision rules (SPSDR)) au problème de couverture en boucle fermée.

Dans SPSDR, nous conservons le peigne de scénarios initial et, en chaque instant, nous effectuons des partitions \mathcal{P}_t de scénarios (de manière identique à la création des nœuds d’un arbre). Chaque paquet de scénarios obtenu est appelé un **ensemble de décisions**. Le principe de cette approche est relativement simple et fondé sur deux hypothèses :

1. puisque la contrainte de raffinement est la cause de l’accroissement exponentiel du nombre de nœuds de l’arbre en programmation stochastique, nous décidons de la supprimer. En particulier, le nombre d’éléments de \mathcal{P}_{t+1} peut être plus faible que celui de \mathcal{P}_t (cf. figure (6.3)).

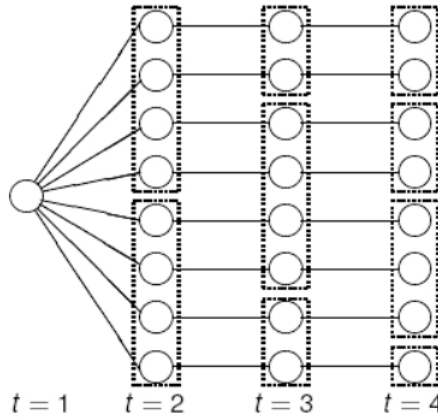


FIG. 6.3 – Construction des ensembles de décisions de l’approche SPSDR

2. nous imposons une unique décision par ensemble de décisions. En fait, nous construisons une règle de décision qui s’adapte à la réalisation de l’aléa.

Nous discuterons dans la section suivante de ces hypothèses et nous détaillerons le fonctionnement de SPSDR.

6.2.3 L'approche SPSDR en détails

SPSDR est donc une alternative à la Programmation Stochastique classique. Les phases de résolution de cette heuristique sont ainsi les mêmes que pour la programmation stochastique, sauf pour la phase de construction des partitions. Comme nous le verrons, cette phase est plus flexible dans SPSDR.

Nous allons présenter les différentes phases de cette approche avant de présenter l'algorithme.

Phase 1 : Construction des ensembles de décisions

Formulation et méthodologie

Comme nous l'avons déjà expliqué, l'approche SPSDR nécessite des partitions \mathcal{P}_t de scénarios en chaque instant, sans obligation de raffinement d'un pas de temps sur l'autre. Chaque élément des partitions est appelé *ensemble de décisions*. L'approche utilisée pour construire ces partitions est inspirée de [7]. Nous définissons une distance entre 2 scénarios (par exemple ℓ_1 , ℓ_2 ou ℓ_∞) sur l'espace des scénarios Ξ . Pour chaque élément de la partition, nous associons un scénario représentatif et nous attribuons une distribution de probabilité à l'ensemble de ces scénarios représentatifs. Le but est de calculer la distance entre 2 distributions : celle de l'ensemble originel des scénarios et celle de l'ensemble des scénarios représentatifs.

Posons brièvement ce problème : soit $\Xi = \{\xi^1, \dots, \xi^N\}$ l'espace des scénarios, μ et $\hat{\mu}$ deux mesures discrètes de probabilité sur cet espace. Nous introduisons les deux probabilités p_i et \hat{p}_i telles que $p_i = \mu\{\xi = \xi^i\}$ et $\hat{p}_i = \hat{\mu}\{\xi = \xi^i\}$.

Nous supposons que le support de $\hat{\mu}$ est $\Xi_J = \{\xi^i\}_{i \in J}$, où $J \subset \{1, \dots, N\}$, impliquant donc que $\hat{p}_i = 0$, $\forall i \notin J$. En se munissant d'une distance $c : \Xi \times \Xi \rightarrow \mathbb{R}$, nous estimons la distance entre les deux mesures de probabilité via la fonction de Kantorovitch ψ :

$$\psi(\mu, \hat{\mu}) = \min_{\eta} \left\{ \sum_{i,j=1}^N c(\xi^i, \xi^j) \eta_{i,j} \mid \eta_{i,j} \geq 0, \forall i, j, \sum_{i=1}^N \eta_{ij} = \hat{p}_j \forall j, \sum_{j=1}^N \eta_{ij} = p_i \forall i \right\}. \quad (6.4)$$

Ce problème (6.4) est un problème linéaire de transport.

Pour construire une partition, il est donc nécessaire de trouver un ensemble d'indices J , tel que $|J| = s$ (s fixé), et une mesure de probabilité $\hat{\mu}$ sur Ξ avec un support Ξ_J telle que la distance $\psi(\mu, \hat{\mu})$ est minimale. Ce problème est équivalent à un problème de p-médianes, affirmation que nous allons prouver.

Preuve: Supposons que J est fixé. Le problème à résoudre est donc :

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

$$\begin{aligned}
 \min_{\hat{p}, \eta} \quad & \sum_{i,j=1}^N c(\xi^i, \xi^j) \eta_{ij} \\
 & \sum_{i=1}^N \eta_{ij} = \hat{p}_j, \quad \forall j = 1, \dots, N \\
 & \sum_{j=1}^N \eta_{ij} = p_i, \quad \forall i = 1, \dots, N \\
 & \eta_{ij} \geq 0, \quad \forall i, j \\
 & \sum_{j=1}^N \hat{p}_j = 1, \quad \hat{p}_j \geq 0, \quad \forall j \in J, \quad \hat{p}_j = 0, \quad \forall j \notin J.
 \end{aligned}$$

La solution optimale de ce problème de transport de masse à J fixé peut être obtenue comme suit. Nous construisons une partition de l'ensemble $I = \{1, \dots, N\}$ en assignant chaque élément i de I à l'élément j de J le plus proche (en cas d'égalité, i est assigné à l'élément ayant le plus petit indice). Les éléments de la partition sont ainsi définis par

$$\Delta(j) = \{i \in I = \{1, \dots, N\} \mid j = \arg \min_{k \in J} c(\xi^i, \xi^k)\}, \quad \forall j \in J.$$

Comme $c(\xi^j, \xi^j) = 0$, alors $j \in \Delta(j)$. Le minimum du problème est ainsi atteint en

$$\hat{p}_j = \begin{cases} \sum_{i \in \Delta(j)} p_i & \text{si } j \in J \\ 0 & \text{sinon} \end{cases}$$

et

$$\eta_{ij} = \begin{cases} 0 & \text{si } j \notin \Delta(J) \\ p_i & \text{sinon.} \end{cases}$$

A la vue de l'analyse précédente, nous concluons que pour trouver l'ensemble J et la mesure de probabilité $\hat{\mu}$, nous devons résoudre un problème d'optimisation combinatoire. En notant $c_{ij} = p_i c(\xi^i, \xi^j)$, le problème d'optimisation devient :

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

$$\min_{x,y} \quad \sum_{i=1}^N \sum_{j=1}^N c_{ij} x_{ij} \quad (6.5)$$

$$\sum_{j=1}^N x_{ij} = 1, \quad \forall i \quad (6.6)$$

$$0 \leq x_{ij} \leq y_j, \quad \forall i, j \quad (6.7)$$

$$\sum_{j=1}^N y_j = s \quad (6.8)$$

$$y_j \in \{0, 1\}, \quad \forall j. \quad (6.9)$$

C'est précisément la formulation du problème de p-médianes. Les médianes du problème sont les j tels que $y_j = 1$. □

Le problème de p-médianes dans notre cas consiste à trouver s scénarios médians parmi les N possibles et d'assigner chacun des $N - s$ scénarios au scénario médian le plus proche. Ce problème est un problème classique en Recherche Opérationnelle et est connu pour être NP-difficile [8]. Il existe des algorithmes pour générer une solution exacte [12] ou pour générer des solutions proches de l'optimal théorique [6].

Nous venons de voir comment créer une partition à un instant donné. Nous pouvons généraliser facilement au cas multipériodique. Les partitions :

- ne sont pas nécessairement des raffinements de l'instant précédent : en conséquence, nous pouvons construire les partitions à chaque instant indépendamment des autres pas de temps ;
- doivent dépendre du passé : pour cela nous considérons le sous-scénario $\hat{\sigma}$ en t qui est une restriction du scénario σ sur la période $\{1, \dots, t\}$. Au lieu de définir une distance à un instant, nous définissons une distance qui prenne en compte toute la période de 1 à t . Il ne reste ensuite qu'à résoudre le problème (6.5) à chaque instant.

Phase 2 : Optimisation des règles de décisions

Nous venons de construire des ensembles de décisions en chaque instant : soit I_t^k le $k^{\text{ème}}$ des $|I_t|$ ensembles de décisions à l'instant t . Nous imposons sur chaque ensemble de décisions une unique décision v_t^k (figure 6.4).

Introduisons la fonction indicatrice $\mathbb{1}_{\{\sigma_t \in I_t^k\}}$ qui vaut 1 si le sous-scénario σ_t est associé à l'ensemble de décisions I_t^k et 0 sinon. Comme, pour un scénario, nous ne pouvons prendre qu'une seule décision par pas de temps, nous pouvons récrire la commande u_t , solution du problème (6.3), sous la forme :

$$u_t(\sigma_t) = \sum_{k=1}^{|I_t|} \mathbb{1}_{\{\sigma_t \in I_t^k\}} v_t^k, \quad \forall t = 2, \dots, T-1. \quad (6.10)$$

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

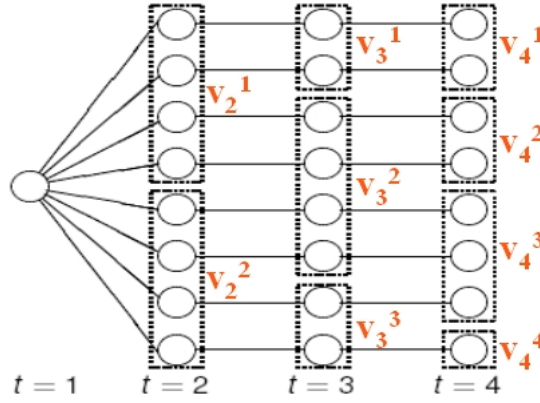


FIG. 6.4 – Règles de décision

Les variables v_t^k sont les véritables décisions du problème d'optimisation. Une fois cette optimisation achevée, pour adapter la décision à un scénario réalisé quelconque σ , il suffit de trouver de quel ensemble de décisions ce scénario appartient (à chaque pas de temps) pour lui appliquer la décision correspondante. Finalement, la décision à prendre n'est pas systématique et dépend de la réalisation de l'aléa : nous avons donc défini une **règle de décision**. u_t étant une somme de fonctions indicatrices, nous avons ainsi défini des **règles de décision constantes par morceaux**.

Phase 3 : Validation

Une fois que la règle de décision finale est calculée, il ne reste plus qu'à évaluer la qualité de cette solution. Pour cela,

1. nous étendons cette règle à l'ensemble des scénarios de Ξ en déterminant pour chaque scénario quelle décision prendre en chaque instant par des calculs de distance. En effet, pour chaque partition construite à chacun des pas de temps t , nous disposons d'un sous-scénario $\sigma_t^s = \{\sigma_1, \dots, \sigma_t\}$ de référence. Pour chaque sous-scénario σ_t , il s'agit donc de trouver quel sous-scénario s de référence est le plus proche de σ_t au sens de la distance c . Ce sous-scénario de référence est donné par $s = \arg \min_{k \leq \nu_t} c(\sigma_t, \sigma_t^k)$. Nous obtenons ainsi un coût de gestion et une évolution de l'état du système différents par scénario ;
2. nous pouvons calculer des indicateurs de comparaison tels que le coût moyen, le nombre de violations de contraintes,...

Utilisation d'experts

Faire une optimisation sur l'ensemble des N scénarios du pool initial peut s'avérer coûteux en temps de calcul. Pour contourner cette difficulté, nous proposons une approche basée sur la

6.2. PROGRAMMATION STOCHASTIQUE AVEC RÈGLES DE DÉCISION CONSTANTES PAR MORCEAUX

création d'**experts**. Un expert n'est rien d'autre qu'un sous-ensemble $\tilde{\Xi}$ de taille $N' \leq N$ de Ξ . Nous le construisons par tirage aléatoire à l'intérieur du pool initial de scénarios. Du point de vue de l'optimisation, nous n'obtenons désormais plus un jeu de commandes u_1, \dots, u_T mais une série de jeux de commandes u_1^l, \dots, u_T^l , $\forall l = 1, \dots, K$ où K est le nombre d'experts :

$$u_t^l(\sigma_t) = \sum_{k=1}^{|I_t^{k,l}|} \mathbb{1}_{\{\sigma_t \in I_t^{k,l}\}} v_t^{k,l}, \quad \forall t = 2, \dots, T-1, \quad \forall l = 1, \dots, K. \quad (6.11)$$

La commande finale qui sera appliquée sera une combinaison convexe des commandes obtenues par chaque expert :

$$u_t(\sigma_t) = \sum_{l=1}^K \alpha_l u_t^l(\sigma_t), \quad \sum_{l=1}^K \alpha_l = 1, \quad \forall t = 1, \dots, T. \quad (6.12)$$

En pratique, nous choisirons $\alpha_l = \frac{1}{K}$, $\forall l$. Au lieu de résoudre un gros problème d'optimisation, nous résolvons désormais K problèmes mais d'une complexité moindre.

Les mêmes types d'approches ont été utilisées dans [9]. De même, la combinaison de conseils d'experts a aussi été exploitée dans le contexte des statistiques et du "machine learning" [3].

Algorithme

L'algorithme associé à cette approche SPSDR est le suivant :

1. - Identifier le processus stochastique et ses paramètres
2. - Générer un grand pool de N scénarios
3. - Répéter pour chaque expert k :
 - (a) - Tirer K_k scénarios parmi les N pour créer l'expert k
 - (b) - Créer les ensembles de décisions sur tous les pas de temps
 - (c) - Optimiser les règles de décisions sur tous les aléas
4. - Créer la règle de décision finale par combinaison convexe des experts
5. - Valider la règle finale sur le pool initial des N scénarios

Parallélisation de SPSDR

Comme nous le verrons en 6.3.3, les temps de calcul peuvent être très élevés, notamment pour créer les partitions. Cependant, SPSDR possède une propriété très intéressante pour sa mise en place : elle peut être massivement parallélisée. En effet, les constructions des partitions à chacun des T pas de temps et les optimisations de chacun des K experts sont indépendants. Nous

pouvons donc faire des gains considérables de temps de calcul en effectuant les KT constructions puis les K optimisations en parallèle.

6.2.4 Résultats obtenus sur le problème de chaînes d'approvisionnement

L'article [14] souligne principalement que, sur un problème de chaîne d'approvisionnement à 12 pas de temps et sur le jeu de données choisi, la programmation stochastique commet une erreur importante de prédiction (de l'ordre de 12%), ce qui représente l'écart entre le coût optimal annoncé après optimisation sur l'arbre et le coût optimal après validation sur un autre échantillon de scénarios. Les performances obtenues en optimisation par SPSDR se retrouvent en simulation puisque l'erreur de prédiction est inférieure à 1%. Enfin, le coût optimal est amélioré par SPSDR de presque 5%. L'optimisation robuste fournit une bonne prise en compte de l'aléa pour les situation extrêmes mais fournit une politique de gestion très coûteuse en comparaison de SPSDR (une augmentation de 10 à 20% du coût selon le niveau de couverture).

6.3 Adaptations de l'approche pour le problème de couverture

Nous ne pouvons en l'état appliquer l'heuristique SPSDR sur le problème de couverture (5.4) du chapitre 5. En effet, il contient des contraintes en probabilité et des aléas multidimensionnels, caractéristiques pour lesquelles SPSDR n'a pas été conçu. Nous allons donc dans cette section détailler les adaptations nécessaires qui ont été effectuées.

6.3.1 Aléa multidimensionnel

Dans le cas unidimensionnel, l'espace des aléas est partitionné en ensembles de décisions. Sur chaque ensemble, une décision unique est imposée. Nous partons de ce principe pour généraliser l'approche au cas multidimensionnel, cas non traité dans [14], mais nécessaire dans le cadre des travaux de cette thèse.

– Aléas non corrélés :

chaque aléa est traité indépendamment. Nous les partitionnons chacun comme dans le cas unidimensionnel, et nous imposons des décisions uniques sur chacun des ensembles de décisions. Autrement dit, cela revient à considérer une partition globale, constituée du produit des partitions de chaque dimension de l'aléa. Désormais, la règle finale est constituée de la combinaison affine des décisions de chaque aléa correspondantes au scénario étudié. Ainsi, la décision prend en compte l'influence de chacun des aléas.

Exemple 6.1 *Illustrons ces propos sur un exemple simple en 2 dimensions. Supposons 2 aléas A et B . Nous partitionnons respectivement A et B en 4 et 3 partitions. Sur chaque partition, notée A_i ou B_i , nous imposons une unique décision notée a_i ou b_i . L'espace global des aléas étant le produit des partitions, la règle décrite s'écrit de la forme :*

	A_1	A_2	A_3	A_4
B_1	$a_1 + b_1$	$a_2 + b_1$	$a_3 + b_1$	$a_4 + b_1$
B_2	$a_1 + b_2$	$a_2 + b_2$	$a_3 + b_2$	$a_4 + b_2$
B_3	$a_1 + b_3$	$a_2 + b_3$	$a_3 + b_3$	$a_4 + b_3$

Quand nous possédons un scénario particulier (\bar{A}, \bar{B}) , il suffit de savoir dans quelles partitions se trouve chaque dimension de l'aléa et d'appliquer la règle correspondante.

En reprenant les notations de la section précédente et en introduisant la variable de décision $v_t^{k,m}$ associée à l'ensemble de décisions $I_t^{k,m}$ défini comme le $k^{\text{ème}}$ ensemble de décisions en t de l'aléa m , nous pouvons récrire la règle de définition sous la forme :

$$u_t(\sigma_t) = \sum_{m=1}^M \sum_{k=1}^{|I_t^m|} \mathbb{1}_{\{\sigma_t \in I_t^{k,m}\}} v_t^{k,m}, \quad \forall t = 2, \dots, T-1. \quad (6.13)$$

– Aléas corrélés :

Dans le cas précédent, les aléas sont supposés indépendants. Mais, par exemple dans notre problème de couverture, les aléas de prix Spot et Week-Ahead sont corrélés. Pour que le principe général de SPSDR demeure simple, une possibilité est de considérer que les aléas corrélés ne constituent qu'un seul aléa. En fait, au moment de construire les partitions, nous définissons une distance qui prend en compte les aléas corrélés entre eux. En reprenant l'exemple précédent et en supposant que A et B sont corrélés, nous définissons une norme du type $c' = \|A - \bar{A}\|_p + \|B - \bar{B}\|_p$. Nous n'obtenons ainsi qu'une seule partition pour les deux aléas A et B , et qui est construite de sorte à considérer l'influence mutuelle des deux. Soit un aléa à M dimensions pouvant être réduit à M' dimensions par effet des corrélations. La règle finale s'écrit donc :

$$u_t(\sigma_t) = \sum_{m=1}^{M'} \sum_{k=1}^{|I_t^m|} \mathbb{1}_{\{\sigma_t \in I_t^{k,m}\}} v_t^{k,m}, \quad \forall t = 2, \dots, T-1. \quad (6.14)$$

Nous pouvons remarquer qu'avec une telle construction des partitions, nous réduisons le nombre de partitions à construire et donc les temps de calcul, le cas multidimensionnel pouvant être coûteux.

Remarque 6.1 Cette gestion du cas multidimensionnel est une idée nouvelle, inédite dans la littérature. Cette règle réduit l'espace des décisions (puisque'il impose une décomposition additive) mais il a comme avantage de ne pas imposer de pondération a priori entre les composantes du processus aléatoire.

Enfin, la nouvelle règle pour chacun des experts doit donc être la somme des règles issues de l'optimisation sur chacune des dimensions m de l'aléa. :

$$u_t^l(\sigma_t) = \sum_{m=1}^M \sum_{k=1}^{|I_t^{m,k,l}|} \mathbb{1}_{\{\sigma_t \in I_t^{k,m,l}\}} v_t^{k,m,l}, \quad \forall t = 2, \dots, T-1, \quad \forall l = 1, \dots, K. \quad (6.15)$$

Seule la définition de la règle a donc été modifiée par l'adaptation au cas multidimensionnel, ce qui n'a aucune conséquence sur l'algorithme général de SPSDR (sous-section 6.2.3).

6.3.2 Contraintes sur l'état

Les contraintes en probabilité du problème de couverture, trop complexes, sont introduites dans le modèle d'optimisation avec un niveau de probabilité de 100%. Ces contraintes peuvent donc être transformées en une contrainte linéaire avec une borne supérieure. Cependant, il peut toujours se réaliser des scénarios sur lesquels la marge ne sera pas suffisante. Donc, pour forcer la stratégie à satisfaire le critère de risque en probabilité lors de la phase de simulation, nous pouvons :

- jouer sur les paramètres du problème : nombre et tailles des experts, nombre d'ensembles de décisions,... Nous pouvons d'ailleurs concevoir que, pour avoir une stratégie plus robuste aux aléas, il faut prendre en compte des experts nombreux, de grande taille avec peu d'ensembles de décisions. Les contraintes seront donc satisfaites sur un plus large espace d'aléas ;
- introduire un coefficient de sécurité sur l'état. Il faut toutefois vérifier qu'une telle technique ne dégrade pas trop la solution.

6.3.3 Comparaisons numériques entre les heuristiques de calcul du problème p -médianes

Pour construire les ensembles de décisions, nous devons résoudre des problèmes de p -médianes. En pratique, nous avons testé deux approches :

1. une heuristique "Variable Neighborhood Decomposition Search" (VNDS) introduite par Mladenovic et Hansen [5]. Cette heuristique est performante, bien que relativement simple. Nous nous dotons de K types de voisinage. En partant d'une solution initiale x , nous générons un point x' aléatoirement dans le premier voisinage de x . A partir de ce nouveau point, nous effectuons une recherche locale de solution pour obtenir une solution x'' . Si cette solution est meilleure que x , nous recommençons cette étape en conservant le même voisinage. Sinon, nous recommençons l'étape à partir du voisinage suivant et cela, jusqu'à ce que tous les voisinages soient parcourus ;
2. une heuristique développée par Vial et Thénier, fondée sur la relaxation lagrangienne du problème de p -médianes [4] et sur une technique d'arrondi aléatoire [11]. Cette relaxation lagrangienne est :

$$\max_{u \geq 0} \left\{ \min_{x, y \in \Omega} \left\{ \sum_{i=1}^N \sum_{j=1}^N c_{ij} x_{ij} + \sum_{j=1}^N \left(1 - \sum_{i=1}^N x_{ij} \right) u_j \right\} \right\}, \quad \Omega = \{x, y : 0 \leq x_{ij} \leq y_j, \forall i, j, \sum_{j=1}^N y_j = s\}.$$

Nous résolvons le dual lagrangien par ACCPM [2]. Cet algorithme permet de reconstruire une solution primale du problème de p -médianes relâché (i.e. sans contrainte d'intégralité). Cette solution fractionnaire est interprétée comme une probabilité qu'un point (scénario)

soit choisi comme médiane (scénario de référence). La sortie de cette étape est une solution fractionnaire β^* telle que $\sum_i \beta_i^* = \nu_t$.

Sur cette base, nous construisons un algorithme aléatoire de tirages successifs de p médianes (p scénarios de référence). Pour un tirage particulier, la partition autour des p scénarios de référence se fait par affectation au plus proche scénario de référence. L'opération est répétée un grand nombre de fois et la meilleure solution est retenue.

La procédure d'arrondi de la solution est la suivante. Pour chaque i nous assignons une variable aléatoire de Bernoulli X_i de probabilité β_i^* . L'élément i est un scénario de référence si $X_i = 1$. Si le nombre de médianes est inférieur à ν_t , un élément i est choisi aléatoirement parmi tous les éléments qui ne sont pas encore scénarios de référence. Un tirage de Bernoulli est calculé. Si la sortie est 1, nous assignons i aux scénarios de référence. Nous itérons ce processus jusqu'à ce que les ν_t scénarios de référence soient calculés.

Pour évaluer ces deux approches, nous avons effectué quelques comparaisons dans le cadre de notre problématique. Nous les avons ainsi testées sur des scénarios de marge de production. Pour plusieurs échantillons de scénarios et pour un certain nombre fixé d'ensembles de décisions, nous comparons dans les tableaux 6.1 et 6.2 :

- les distances cumulées optimales obtenues en chaque instant par les deux approches. Pour chaque scénario, nous calculons la distance à son scénario de référence et nous sommions toutes ces distances. Moins ce cumul est élevé, meilleure est la répartition ;
- le temps moyen de construction des partitions pour un pas de temps.

Quand l'ensemble de scénarios considéré ou le nombre d'ensembles de décisions est peu élevé, les résultats entre les deux approches sont proches voire même identiques. En revanche, quand la taille (et donc la difficulté) croît, des différences apparaissent en faveur de VNS, mais demeurent raisonnables. Si la qualité de la solution est le critère principal de choix, l'heuristique VNS est dans ce cas le choix adéquat. Cependant, un aspect important est la nécessité de disposer d'un outil qui soit capable d'effectuer des calculs en un temps raisonnable, quitte à perdre un peu en qualité. Or, VNS est assez gourmand de ce point de vue : les calculs vont de 300 s à 1300 s pour effectuer une partition en un instant. Or, nous devons faire ce calcul pour $T - 2$ pas de temps (un unique ensemble de décisions en $t = 1$ et aucune décision prise en $t = T$) et pour chaque aléa considéré. Pour illustrer ceci sur notre problème, les temps de calculs peuvent ainsi varier de $8 * 3 * 300s = 2H$ à $8 * 3 * 1300s = 8H40mn$. Et cela pour un seul expert ... En comparaison, les temps obtenus par l'approche basée sur la relaxation lagrangienne sont bien moins élevés (de $5mn$ à $4H20mn$). Mais, nous pouvons noter que si cette technique est beaucoup plus efficace quand le nombre de scénarios est faible, le rapport n'est plus que de un à deux avec 1000 scénarios. Cependant, les gains sont tels que, en pratique, nous préférons cette dernière heuristique.

6.4. FORMULATION SIMPLIFIÉE DU PROBLÈME

Données (V=[10;10;12;12;14;14;16;16])		Fonction objectif du problème p -médianes (R. Lag.)								Durée (R. Lag.)
Nbre Scénarios	Médianes	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	Temps
100	.5xV	74.4	319.1	505.6	742.2	899.9	1138.6	1364.9	1649.5	13s
100	1xV	33.4	216.3	377.6	564.4	716.5	923.1	1091.7	1299.2	20s
100	2xV	13.1	132.5	253.2	398.8	498.4	650.9	747.4	898.9	21s
100	3xV	7.5	90.4	175.4	292.1	359.3	475.7	507.6	621.4	21s
500	.5xV	374.9	1675.9	2512.6	3715.9	4613.5	5789.5	6980.1	8244.9	140s
500	1xV	204.3	1245.4	2055.4	3016.2	3846.6	4975.3	6052.7	7314.6	130s
500	2xV	108.7	873.8	1580.6	2476.2	3214.8	4189.6	5066.4	6223.5	155s
500	3xV	70.3	690.7	1326.6	2198.2	2883.3	3773.7	4611.7	5591.9	163s
1000	.5xV	785.3	3369.9	5195.6	7589.7	9461.2	12034.5	14303.8	17230.2	461s
1000	1xV	412.2	2567.1	4324.7	6347.5	8002.5	10334.9	12479.8	15169.3	468s
1000	2xV	228.6	1818.9	3435.4	5375.5	6867.5	8787.5	10918.3	13236.3	629s
1000	3xV	163.3	1518.5	3044.1	4818.4	6179.3	8052.4	10060.8	12183.6	651s

TAB. 6.1 – Heuristique p -médianes basée sur la relaxation lagrangienne du problème p -médianes

Données (V=[10;10;12;12;14;14;16;16])		Fonction objectif du problème p -médianes (VNS)								Durée (VNS)
Nbre Scénarios	Médianes	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	Temps
100	.5xV	74.4	319.1	505.6	742.2	899.9	1138.6	1364.9	1649.5	302s
100	1xV	33.4	216.3	376.1	564.4	716.5	922.7	1091.7	1299.2	307s
100	2xV	13.1	132.4	252.9	398.8	498.4	650.9	747.4	898.9	306s
100	3xV	7.5	90.4	175.4	292.1	359.3	475.7	507.6	621.4	307s
500	.5xV	379.9	1675.9	2512.6	3697.8	4581.9	5771.2	6952.2	8226.8	544s
500	1xV	204.9	1220.2	2011.6	2984.4	3795.1	4899.1	5963.3	7145.1	575s
500	2xV	96.4	844.8	1528.2	2425.8	3161.1	4121.3	5026.9	6098.2	492s
500	3xV	61.8	674.3	1295.6	2118.2	2804.1	3712.2	4534.5	5498.7	540s
1000	.5xV	785.3	3369.9	5172.2	7589.7	9443.5	11927.9	14275.1	17030.7	582s
1000	1xV	412.2	2492.8	4152.9	6208.3	7882.2	10093.2	12263.9	14724.5	687s
1000	2xV	202.5	1762.9	3271.7	5126.3	6627.1	8585.6	10594.1	12818.1	1260s
1000	3xV	133.9	1429.8	2835.4	4582.3	5986.9	7815.3	9703.8	11772.6	1330s

TAB. 6.2 – Heuristique p -médianes VNS

Les tableaux précédents montrent des temps de calculs assez importants. Une parallélisation de cette construction par pas de temps pourrait permettre dans ce cas une division des temps de calculs par 8. Cette opération se justifie d'autant plus que ces calculs doivent être effectués pour chacune des dimensions de l'aléa et pour chaque expert. En considérant 10 experts et 3 aléas, nous pouvons réduire théoriquement les temps de calculs consacrés à la construction des partitions par 240 !

6.4 Formulation simplifiée du problème

Dans cette partie exploratoire sur l'étude de la boucle fermée, nous avons décidé de traiter un problème simplifié, éliminant quelques difficultés mais permettant tout de même de tirer de premières conclusions.

6.4. FORMULATION SIMPLIFIÉE DU PROBLÈME

Les deux simplifications consistent à :

1. considérer un **problème simple de petite taille** : nous considérons un marché SPOT, un unique contrat Week-Ahead (WA) et seulement deux semaines d'étude de chacune 5 jours. Nous ne considérons donc pas les Week-End Ahead ;
2. faire une **agrégation des différents contrats d'options tarifaires (OT) en un unique stock** : cette simplification est vraiment une approximation qui peut conduire à des comportements différents mais l'objectif est, dans un premier temps, de supprimer les variables binaires du problème provenant de la gestion de ces contrats (déclenchement ou non).

En considérant un horizon de T jours, divisés en 2 semaines W_1 et W_2 , introduisons les notations nécessaires à l'écriture du problème :

Nom	Intervalle de définition	Signification
x_t	$t = 1, \dots, T - 1$	Quantité d'électricité achetée (ou vendue) en t et livrée en $t + 1$ sur le marché SPOT
y_t	$t = 1, \dots, \frac{T}{2}$	Quantité d'électricité achetée (ou vendue) en t et livrée identiquement en chaque jour de W_2 en Week Ahead (WA)
z_t	$t = 1, \dots, T - 1$	Quantité d'électricité déclenchée en t par option tarifaire pour effacement en $t + 1$
C_t^S	$t = \frac{T}{2}, \dots, T - 1$	Prévision de prix du Spot en t
C_t^W	$t = 1, \dots, T - 1$	Prévision de prix du WA en t
C_t^{OT}	$t = 1, \dots, T - 1$	Prix de déclenchement d'option tarifaire en t
M_t	$t = 1, \dots, T$	Marge physique du parc en EDF en t
S_t	$t = 1, \dots, T$	Stock d'électricité disponible en OT en t

Ce problème s'écrit donc :

$$\begin{aligned}
 \min_{x,y,z} \quad & \mathbb{E} \left[\frac{T}{2} \sum_{t=1}^{T/2} C_t^W y_t + \sum_{t=1}^{T-1} C_t^S x_t \right] + \sum_{t=1}^{T-1} C_t^{OT} z_t \\
 \text{s.t.} \quad & \begin{cases} \mathbb{P}[M_t + x_{t-1} + (\sum_{\tau=1}^{T/2} y_\tau) \cdot \mathbb{1}_{\{t \in W_2\}} + z_{t-1} \geq \mathcal{M}_S] \geq 1 - p, \forall t \in 2, \dots, T \\ S_T = S_1 - \sum_{\tau=1}^{T-1} z_\tau \geq 0 \\ \underline{x}_t \leq x_t \leq \bar{x}_t, \forall t = 1, \dots, T - 1 \\ \underline{y}_t \leq y_t \leq \bar{y}_t, \forall t = 1, \dots, \frac{T}{2} \\ 0 \leq z_t \leq \bar{z}_t, \forall t = 1, \dots, T - 1, \end{cases}
 \end{aligned} \tag{6.16}$$

où $\mathbb{1}_{\{X\}}$ est la fonction indicatrice valant 1 si X est vraie et 0 sinon, et p le niveau de défaillance tolérée.

Ce problème demeure un problème linéaire sous contrainte en probabilité. Cependant la taille du problème est particulièrement petite puisqu'il n'est constitué que de $\frac{5T}{2} - 2$ variables et $\frac{7T}{2} - 2$ contraintes. Dans notre exemple, cela correspond à 23 variables et 33 contraintes.

Remarque 6.2 *En pratique, nous pouvons remplacer le terme $(\sum_{\tau=1}^{T/2} y_{\tau}) \cdot \mathbb{1}_{\{t \in W_2\}}$ par la somme plus détaillée $\sum_{\tau=(\lceil \frac{t}{T/2} \rceil - 2) \cdot \frac{T}{2} + 1}^{(\lceil \frac{t}{T/2} \rceil - 1) \cdot \frac{T}{2}} y_{\tau}$. Cependant, pour des raisons de lisibilité, nous conserverons la première écriture.*

6.5 Ecriture du problème de couverture avec SPSDR

Pour formuler le problème avec l'approche SPSDR, récapitulons les notations nécessaires en plus des notations de la sous-section 6.4 :

Nom	Intervalle de définition		Signification
M			Nombre d'incertitudes
N			Nombre de scénarios
σ_t^n	$t = 1, \dots, T - 1$	$n = 1, \dots, N$	Restriction du scénario n à la période $\{1, \dots, t\}$
$ I_t^m $	$t = 1, \dots, T - 1$	$m = 1, \dots, M$	Nombre d'ensembles de décisions créés au pas de temps t pour l'incertitude m
$I_t^{k,m}$	$t = 1, \dots, T - 1$ $k = 1, \dots, I_t^m $	$m = 1, \dots, M$	$k^{\text{ème}}$ ensemble de décisions de la période t pour l'incertitude m
$\mathbb{1}_{\{\sigma_t^n \in I_t^{k,m}\}}$	$t = 1, \dots, T - 1$ $k = 1, \dots, I_t^m $	$m = 1, \dots, M$ $n = 1, \dots, N$	Fonction indicatrice valant 1 si le sous-scénario σ_t^n appartient à l'ensemble de décisions $I_t^{k,m}$ et 0 sinon
$\hat{x}_t^{k,m}, \hat{y}_t^{k,m}, \hat{z}_t^{k,m}$	$t = 1, \dots, T - 1$ $k = 1, \dots, I_t^m $	$m = 1, \dots, M$ $n = 1, \dots, N$	Variable de SPOT, de WA et d'OT par ensemble de décisions et par pas de temps
x_t^n, y_t^n, z_t^n	$t = 1, \dots, T - 1$	$n = 1, \dots, N$	Variable finale de SPOT de WA et d'OT par pas de temps
$M_{t,n}, C_{t,n}^S, C_{t,n}^W$	$t = 1, \dots, T$	$n = 1, \dots, N$	$n^{\text{ème}}$ scénario de marge, coûts SPOT et WA en t

La formulation ci-dessous est la formulation pour un unique expert. Pour K experts, il suffit de résoudre K fois ce problème et de moyenner les solutions obtenues.

(6.17)

Nous générons 75000 scénarios de marge, de prix Spot et de Week-ahead.

Nous limiterons les achats Spot à 6000 MW par jour et les ventes à 2000 MW uniquement en première semaine (nous considérons que la seconde semaine n'a qu'un rôle de couverture à terme).

Nous limitons chaque décision d'achat Week-ahead à 500 MW par jour (pour livraison sur les 5 jours de la semaine suivante soit 2500 MW). Les ventes sont limitées à 220 MW.

Nous faisons également l'hypothèse que les prix de vente sont toujours plus bas que les prix d'achat (pour limiter les arbitrages et pour modéliser le fait que, généralement, EDF est un acteur important du marché entraînant des hausses de prix quand EDF est court¹ et des baisses de prix quand EDF est long).

Nous considérons également un stock d'options tarifaires de 2000 MW. Nous limitons les effacements à 500 MW par jour moyennant 100 euros/MW.

Enfin, la marge de sécurité est fixée à 3000 MW par jour et nous souhaitons nous couvrir contre le risque que la marge soit inférieure à cette marge de sécurité dans plus de 1% des scénarios possibles.

6.6.2 Modèles simples de comparaison

Pour évaluer les résultats de l'approche SPSDR, nous présentons deux modèles simples de comparaison. Ces modèles sont anticipatifs puisque nous prenons des décisions en t_0 sur les deux semaines. Une approche glissante pourrait corriger cela mais nécessiterait des générations de scénario de marge après chaque jour. Or, rappelons que les scénarios de marge sont calculés en quelques heures de temps de calcul...

Déterministe avec contrainte en probabilité

La première approche est une approche déterministe avec contrainte en probabilité. Il s'agit en fait d'appliquer le modèle (9.6) du chapitre 5. Nous calculons le quantile à 1% et nous optimisons le problème linéaire de couverture sur la base des coûts moyens. Nous simulons ensuite la solution obtenue sur les 75000 scénarios. Le tableau 6.5 montre les décisions prises en fonction des quantiles calculés :

¹En manque d'électricité. Dans le cas contraire, EDF est long.

6.6. RÉSULTATS NUMÉRIQUES : PRÉ-REQUIS

	Jour 2	Jour 3	Jour 4	Jour 5	Jour 6	Jour 7	Jour 8	Jour 9	Jour 10
Quantile 99%	2914	6324	6296	−1881	845	349	−3019	−2640	2213
Spot	0	−2000	−2000	4881	2636	3132	6000	6000	1268
Week-Ahead	0	0	0	0	−481	−481	−481	−481	−481
Options Tarifaires	86	0	0	0	0	0	500	121	0
Bilan	3000	4324	4296	3000	3000	3000	3000	3000	3000

TAB. 6.3 – Solution du problème de couverture contre le risque 1%

Nous appliquons ces commandes sur les 75000 scénarios. Nous comptabilisons le nombre de scénarios sur lesquels la marge n'est pas supérieure à 3000 MW (Tab. 6.4).

	Jour 2	Jour 3	Jour 4	Jour 5	Jour 6	Jour 7	Jour 8	Jour 9	Jour 10	Coût moyen
Violations	0	0	0	0	0	0	0	0	0	1090858 €

TAB. 6.4 – Pourcentage de violations et coût moyen en simulation de la contrainte de respect du risque 1% de la contrainte de marge sur les 75000 scénarios

Synthèse 6.1 (Modèle déterministe sous contrainte en probabilité) *Cette approche permet de satisfaire le risque 1% puisque les violations en chaque jour sont nulles (Tab. 6.4). Le coût moyen de cette couverture est de 1,09.10⁶ euros. Cette approche semble sur-protectrice puisque les niveaux de violation n'approchent pas les 1% autorisés par jour.*

Robustesse

Dans le problème (9.6), les seules contraintes que nous devons rendre robustes sont les $T - 1$ contraintes en probabilité et cela par rapport à l'aléa de marge de production. A l'aide de méthodologies déjà précisées dans la partie précédente, nous cherchons donc à approximer les contraintes en probabilité

$$\mathbb{P}[x_{t-1} + \bar{M}_t + \sigma_t \epsilon_t + (\sum_{\tau=1}^{T/2} y_\tau) \cdot \mathbb{1}_{\{t \in W_2\}} + z_{t-1} \geq \mathcal{M}_S] \geq 1 - p, \quad \forall t \in 2, \dots, T$$

par les contraintes linéaires

$$x_{t-1} + \bar{M}_t + \sigma_t \epsilon_t + (\sum_{\tau=1}^{T/2} y_\tau) \cdot \mathbb{1}_{\{t \in W_2\}} + z_{t-1} \geq \mathcal{M}_S, \quad \forall \epsilon_t \in \Xi_t, \quad \forall t \in 2, \dots, T$$

où \bar{M}_t étant la marge moyenne, σ_t l'écart type de la marge et Ξ_t est un ensemble d'incertitude construit de sorte à se prémunir avec une probabilité $1 - p$.

6.6. RÉSULTATS NUMÉRIQUES : PRÉ-REQUIS

Pour chaque contrainte, un seul aléa ϵ_t intervient. Nous pourrions, comme dans les chapitres 3 et 4, utiliser une norme $\ell_1 \cap \ell_\infty$. Or, dans notre cas, nous avons $\|\epsilon_t\|_1 = \|\epsilon_t\|_\infty = |\epsilon_t|$, et il suffit donc de résoudre le problème $\min_{\epsilon_t} \{\sigma_t \epsilon_t, |\epsilon_t| \leq \kappa\}$. La solution vaut $-\kappa|\sigma_t| = -\kappa\sigma_t$ (σ_t étant positif par définition en tant qu'écart-type).

L'équivalent robuste dans ce cas-ci est donc :

$$\begin{aligned} \min_{x,y,z} \quad & \frac{T}{2} \sum_{t=1}^{T/2} \bar{C}_t^W y_t + \sum_{t=1}^{T-1} \bar{C}_t^S x_t + \sum_{t=1}^{T-1} C_t^{OT} z_t \\ \text{s.t.} \quad & \begin{cases} \bar{M}_t + x_{t-1} - \kappa\sigma_t + \left(\sum_{\tau=1}^{T/2} y_\tau\right) \cdot \mathbb{1}_{\{t \in W_2\}} + z_{t-1} \geq \mathcal{M}_S, \quad \forall t \in 2, \dots, T \\ S_T = S_1 - \sum_{\tau=1}^{T-1} z_\tau \geq 0 \\ \underline{x}_t \leq x_t \leq \bar{x}_t, \quad \forall t = 1, \dots, T-1 \\ \underline{y}_t \leq y_t \leq \bar{y}_t, \quad \forall t = 1, \dots, \frac{T}{2} \\ 0 \leq z_t \leq \bar{z}_t, \quad \forall t = 1, \dots, T-1. \end{cases} \end{aligned} \quad (6.18)$$

Appliquons cette approche sur le jeu d'essais 6.6.1. Voici un tableau récapitulatif de résultats avec une couverture à 99% :

	Jour 2	Jour 3	Jour 4	Jour 5	Jour 6	Jour 7	Jour 8	Jour 9	Jour 10
Moyenne	3607	7303	7597	-318,5	2042	1524	-1522	-916	3898
Spot	-607	-2000	-2000	3318	-1542	0	2022	1416	0
Week-Ahead	0	0	0	0	2500	2500	2500	2500	2500
Options Tarifaires	0	0	0	0	0	0	0	0	0
Bilan	3000	5303	5597	3000	3000	4024	3000	3000	6398

TAB. 6.5 – Solution du problème de couverture contre le risque 1%

	Jour 2	Jour 3	Jour 4	Jour 5	Jour 6	Jour 7	Jour 8	Jour 9	Jour 10	Coût moyen
Violations	0,95	0	0	0,9	0,95	0,37	0,57	0,17	0,01	-396 094 €

TAB. 6.6 – Pourcentage de violations et coût moyen en simulation de la contrainte de respect du risque 1% de la contrainte de marge sur les 75000 scénarios

Synthèse 6.2 (Modèle robuste) *Contrairement à l'approche déterministe, un certain pourcentage de violations apparaît en chaque pas de temps (Tab. 9.4). Cependant, le niveau maximal de 1% par jour est respecté. Si dans le modèle déterministe les coûts de couverture sont positifs, le modèle robuste permet de dégager des bénéfices (en partie car il s'est permis plus de violations).*

Paramétrages empiriques

Nous avons vu dans les deux sous-sections précédentes qu'un niveau de couverture fixé par avance conduit, en simulation, à une sur-protection. Pour améliorer le comportement des modèles et surtout les coûts, nous pouvons ajuster empiriquement les niveaux de couverture afin d'atteindre les 1% de violations par pas de temps.

Le premier constat est qu'un bon paramétrage permet au modèle déterministe avec contrainte en probabilité de passer d'un coût de $1,09.10^6$ € à un gain de $6,4.10^5$ €. Cet écart est considérable et montre que ce modèle est très sensible au paramétrage et donc difficile à utiliser. Ceci peut s'expliquer, car implicitement, nous approximons la marge par une loi normale. Or, cette approximation n'est pas toujours satisfaisante (voir Annexe B). Un choix de niveau de couverture trop élevé peut ainsi coûter très cher et doit donc être évité avec ce modèle.

Avec un pourcentage de violations quasi-identique et atteignant le maximum autorisé en chaque instant, les bénéfices obtenus à l'aide des deux approches sont de $6,4.10^5$ €. Cependant, ce coût est une **borne maximale** sur les gains potentiels. En effet, pour obtenir ces gains, nous devons modifier les niveaux de couverture par tâtonnement en simulation. Or, en pratique, il serait impossible de faire des tâtonnements aussi précis, ce qui conduirait donc à des bénéfices moindres.

	Coût moyen (euros)
Déterministe	-639 703
Robuste	-641 489

TAB. 6.7 – Coûts moyens après ajustement empirique des niveaux de couverture (violations à 1% par pas de temps)

Synthèse 6.3 (Paramétrages empiriques) *Avec les deux approches, la borne maximale sur les bénéfices est de $6,4.10^5$ €.*

6.7 Résultats numériques avec SPSDR

Dans les résultats suivants, nous utilisons la méthodologie présentée dans ce chapitre. Pour gérer le cas multidimensionnel de l'aléa, nous couplons les aléas de prix (Spot et Week-Ahead) (section 6.2.3) en construisant un unique ensemble de décisions prenant en compte ces deux aléas conjointement. L'aléa de marge est lui considéré indépendamment. Cela signifie que la règle de décision finale possède deux composantes : une pour les aléas conjoints Spot et Week-Ahead et une pour la marge.

Dans cette section, nous allons étudier l'effet de trois paramètres inhérents à l'heuristique SPSDR :

1. le nombre d'ensembles de décisions ;

2. la taille des experts ;
3. le nombre d'experts.

L'objectif de cette section est de voir si cette approche permet :

- de satisfaire le critère de risque 1% ;
- de dégager des bénéfices.

6.7.1 Effets du nombre d'ensembles de décisions et de la taille des experts

Nous fixons le nombre d'experts à 1.

Ensembles de décisions croissants

Dans le tableau 6.8 suivant, nous pouvons dégager les informations suivantes :

- A taille d'expert fixée, un nombre d'ensembles de décisions plus important conduit à des violations de plus en plus importantes. Cela s'explique par le fait que, si le nombre d'ensembles est faible, le nombre de scénarios associés à chaque ensemble est plus important. Chaque décision est ainsi plus robuste puisque basée sur plus d'informations ;
- A nombre d'ensembles de décisions fixé, plus la taille d'expert est importante moins les violations sont importantes. L'explication est la même que dans le cas précédent : plus le nombre de scénarios est important, plus il y a de scénarios dans chaque ensemble de décisions. La décision est ainsi plus robuste ;
- Il existe un jeu de paramètres qui permet de satisfaire le niveau de violation maximum de 1%.

1 expert	
Taille Experts	5 5 6 6 7 7 8 8
50	0.67 0 0 1.15 4.70 2.22 8.59 7.38 0.17
100	1.05 0 0 1.85 2.67 2.28 5.37 5.75 0.17
200	0.15 0 0 0.56 2.78 1.51 4.97 4.24 0.17
500	0.06 0 0 0.59 2.61 1.59 2.33 2.59 0.17
1000	0.03 0 0 0.19 1.26 1.05 1.13 1.36 0.12
2000	0.05 0 0 0.16 0.65 0.38 0.54 0.72 0.15
10 10 12 12 14 14 16 16	
50	1.05 0 0 2.43 5.87 3.49 8.35 8.43 0.17
100	0.14 0 0 1.69 4.96 2.69 8.21 6.29 0.17
200	0.04 0 0 1.81 5.12 3.31 4.74 4.97 0.17
500	0.17 0 0 0.87 2.69 1.83 3.44 3.14 0.17
1000	0.01 0 0 0.39 1.48 1.21 2.56 2.89 0.12
2000	0.01 0 0 0.37 0.83 0.92 1.13 1.17 0.14
20 20 24 24 28 28 32 32	
50	1.05 0 0 1.54 5.73 3.38 10.4 14.4 0.16
100	1.05 0 0 1.46 4.56 3.45 8.79 6.97 0.17
200	0.08 0 0 1.36 4.66 3.21 8.17 6.65 0.15
500	0.23 0 0 0.88 3.57 2.82 4.79 4.53 0.17
1000	0.01 0 0 0.76 2.59 1.86 3.43 3.32 0.17
2000	0.01 0 0 0.47 1.52 1.28 2.61 2.50 0.09

TAB. 6.8 – Pourcentage de violations de la contrainte de marge obtenues en simulation avec l'heuristique SPSDR

6.7. RÉSULTATS NUMÉRIQUES AVEC SPSDR

Le tableau 6.9 représente les coûts moyens associés aux niveaux de violation du tableau 6.8. Le seul coût représentatif est le coût où le critère 1% est satisfait et est de -346791 €. Ces bénéfices sont inférieurs de moitié à la borne maximale du tableau 6.7. Ce bénéfice est aussi inférieur à celui obtenu par l'approche robuste avec 99% de couverture.

L'approche SPSDR permet donc, avec un certain paramétrage, de satisfaire les objectifs (respect de la contrainte et obtention d'un bénéfice). Cependant, les bénéfices ne sont pas satisfaisants.

		1 expert															
Taille	Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
50		-1	107	942						-1	170	376					
100		-980	663							-937	062						
200		-821	973							-871	348						
500		-685	051							-805	387						
1000		-445	942							-649	509						
2000		-346	791							-491	359						

TAB. 6.9 – Coût moyen de simulation obtenu avec l'heuristique SPSDR

Ensembles de décisions constants

Dans cette sous-section, nous fixons les ensembles de décisions constants dans le temps. Avec cette situation, nous contredisons complètement le principe de la programmation stochastique classique qui impose un accroissement de l'arbre dans le temps.

		1 expert									
Taille	Experts	5 ensembles de décisions par jour									
500		0.01	0	0	0.48	2.34	1.38	1.15	1.84	0.17	
1000		0.06	0	0	0.13	0.66	0.49	1.23	0.81	0.15	
Taille	Experts	10 ensembles de décisions par jour									
500		0.15	0	0	0.29	1.80	2.18	3.49	2.80	0.09	
1000		0.05	0	0	0.45	1.62	1.14	2.16	1.78	0.16	
Taille	Experts	20 ensembles de décisions par jour									
500		0.37	0	0	0.39	2.45	1.95	4.28	3.91	0.16	
1000		0.07	0	0	0.43	1.98	1.22	4.04	2.87	0.12	

TAB. 6.10 – Pourcentage de violations de la contrainte de marge avec une approche SPSDR et un nombre d'ensembles de décisions constants dans le temps

Si nous mettons en parallèle les résultats ci-dessus avec ceux du tableau 6.8, nous nous apercevons que le fait que les ensembles de décisions soient constants ne joue pas réellement. Les niveaux de violation diminuent mais surtout par le fait que le nombre d'ensembles de décisions est plus faible. La solution est ainsi plus robuste. Cela montre que le nombre d'ensembles de décisions par jour doit être essentiellement choisi en fonction de la difficulté, en ce pas de temps, à satisfaire le niveau de violations. Si le nombre de violations risque d'être important, nous choisissons un

6.7. RÉSULTATS NUMÉRIQUES AVEC SPSDR

faible nombre d'ensembles pour avoir une décision robuste. Sinon, nous essayons de prendre un nombre suffisant pour que la règle soit la plus adaptative possible.

6.7.2 Effets du nombre d'experts

1 expert																		
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16		
50	0.67	0	0	1.15	4.70	2.22	8.59	7.38	0.17	1.05	0	0	2.43	5.87	3.49	8.35	8.43	0.17
100	1.05	0	0	1.85	2.67	2.28	5.37	5.75	0.17	0.14	0	0	1.69	4.96	2.69	8.21	6.29	0.17
200	0.15	0	0	0.56	2.78	1.51	4.97	4.24	0.17	0.04	0	0	1.81	5.12	3.31	4.74	4.97	0.17
500	0.06	0	0	0.59	2.61	1.59	2.33	2.59	0.17	0.17	0	0	0.87	2.69	1.83	3.44	3.14	0.17
1000	0.03	0	0	0.19	1.26	1.05	1.13	1.36	0.12	0.01	0	0	0.39	1.48	1.21	2.56	2.89	0.12
2000	0.05	0	0	0.16	0.65	0.38	0.54	0.72	0.15	0.01	0	0	0.37	0.83	0.92	1.13	1.17	0.14
5 experts																		
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16		
50	0.45	0	0	1.74	4.13	3.19	8.04	6.48	0.17	0.43	0	0	1.80	4.92	3.27	8.79	7.41	0.17
100	0.35	0	0	0.99	3.01	2.37	5.87	5.27	0.17	0.43	0	0	1.76	3.36	3.12	7.24	6.09	0.17
200	0.29	0	0	0.82	2.65	2.04	3.06	3.41	0.16	0.37	0	0	0.68	3.57	2.25	4.40	4.45	0.16
500	0.17	0	0	0.29	1.39	1.25	1.67	1.71	0.17	0.25	0	0	0.52	1.77	1.76	2.82	2.34	0.16
1000	0.09	0	0	0.37	0.68	0.69	1.10	0.74	0.16	0.07	0	0	0.35	1.04	1.15	1.59	1.91	0.15
2000	0.04	0	0	0.18	0.37	0.35	0.34	0.61	0.14	0.04	0	0	0.28	0.73	0.71	0.78	0.99	0.15
10 experts																		
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16		
50	0.67	0	0	1.98	4.99	3.47	7.67	7.01	0.17	0.99	0	0	2.54	5.37	3.03	7.75	7.83	0.17
100	0.45	0	0	1.37	3.27	2.11	5.26	5.33	0.17	0.51	0	0	1.57	3.98	2.88	6.82	6.21	0.17
200	0.36	0	0	0.63	1.88	1.79	3.59	3.31	0.16	0.43	0	0	0.76	2.98	2.26	4.81	4.78	0.17
500	0.22	0	0	0.41	1.44	1.14	1.91	1.52	0.16	0.16	0	0	0.63	1.96	1.41	2.26	2.83	0.16
1000	0.13	0	0	0.18	0.72	0.68	0.88	0.91	0.16	0.07	0	0	0.27	1.06	1.05	1.35	1.42	0.16
2000	0.06	0	0	0.12	0.34	0.37	0.39	0.48	0.13	0.06	0	0	0.23	0.61	0.66	0.69	0.76	0.14

TAB. 6.11 – Pourcentages de violation du critère de risque 1% avec l'approche SPSDR : effet du nombre d'experts

Nous avons vu précédemment l'effet de la taille des experts ainsi que du nombre d'ensembles de décisions. Dans cette sous-section, nous pouvons nous rendre compte de l'importance de multiplier le nombre d'experts. Nous voyons que plus le nombre d'experts est important, plus la solution devient robuste puisque les niveaux de violations diminuent, à nombre d'ensembles de décisions et à taille d'expert fixés. Pour exemple, pour 1 expert, le critère 1% n'est pas satisfait dans le cas d'ensembles de décisions fixés à 10, 10, 12, 12, 14, 14, 16, 16. Mais, juste en augmentant le nombre d'experts, ce critère est satisfait pour 5 et 10 experts. En effet, plus le nombre d'experts est important, plus le nombre de scénarios pris en compte pour l'optimisation des règles est important. Avec 1 expert, la règle est construite avec l'aide de, par exemple, 1000 scénarios. Avec 10 experts, il y a potentiellement 10 000 scénarios qui peuvent apporter de l'information à l'optimisation de la règle, ce qui rend la règle nécessairement plus riche et mieux adaptée.

6.7. RÉSULTATS NUMÉRIQUES AVEC SPSDR

1 expert																
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
50	-1	107	942								-1	170	376			
100		-980	663								-937	062				
200		-821	973								-871	348				
500		-685	051								-805	387				
1000		-445	942								-649	509				
2000		-346	791								-491	359				
5 experts																
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
50		-1	093	476							-1	121	652			
100		-985	317								-1	054	842			
200		-883	811								-959	714				
500		-665	324								-829	663				
1000		-526	201								-674	019				
2000		-381	066								-528	946				
10 experts																
Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
50		-1	126	701							-1	168	054			
100		-1	001	632							-1	073	026			
200		-870	201								-986	221				
500		-714	686								-785	455				
1000		-529	283								-642	826				
2000		-390	475								-531	785				

TAB. 6.12 – Coûts moyens obtenus en simulation avec l’approche SPSDR : effet du nombre d’experts

Au niveau des coûts, une règle plus riche permet de dégager des bénéfices bien plus importants. Nous voyons que nous pouvons obtenir des bénéfices de l’ordre de 346 791 € avec un unique expert et de 531 785 € à l’aide de 10 experts. Ce gain est bien supérieur à ce que les approches statiques permettaient d’obtenir et est très peu inférieur à la borne maximale (641 489 €). Cette approche en boucle fermée s’adapte ainsi à la réalisation des aléas et est plus performante que les approches statiques.

6.7.3 Introduction d’un coefficient d’immunité

L’objectif principal est la satisfaction du niveau maximal de violation à 1%. Or, avec SPSDR, aucun résultat ne peut nous assurer que ce critère sera satisfait en simulation. Pour cela, nous pouvons introduire un coefficient d’immunité κ dans l’optimisation afin de rajouter une nouvelle sécurité au sein du modèle. La contrainte prise en compte est :

$$x_{t-1} + M_t + \left(\sum_{\tau=1}^{T/2} y_{\tau} \right) \cdot \mathbb{1}_{\{t \in W_2\}} + z_{t-1} \geq \kappa \cdot \mathcal{M}_S, \quad \forall \epsilon_t \in \Xi_t, \quad \forall t \in 2, \dots, T.$$

Il s’agit en fait d’une approche robuste simple.

6.7. RÉSULTATS NUMÉRIQUES AVEC SPSDR

1 expert														
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14
1	500	0.06	0	0	0.59	2.61	1.59	2.33	2.59	0.17	0.17	0	0	0.87
1.1	500	0.03	0	0	0.66	1.15	1.17	0.95	1.88	0.17	0	0	0.63	1.93
1.2	500	0	0	0	0.37	1.10	0.72	1.42	1.49	0.15	0.04	0	0	1.11
1.3	500	0	0	0	0.17	0.57	0.62	1.05	1.38	0.09	0	0	0	0.34
1	1000	0.03	0	0	0.19	1.26	1.05	1.13	1.36	0.12	0.01	0	0	0.39
1.1	1000	0.40	0	0	0.26	0.78	0.60	0.70	0.75	0.09	0.04	0	0	0.43
1.2	1000	0	0	0	0.14	0.34	0.48	0.46	0.69	0.09	0	0	0	0.23
1.3	1000	0	0	0	0.16	0.31	0.17	0.09	0.42	0.08	0.04	0	0	0.21
5 experts														
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14
1	500	0.17	0	0	0.29	1.39	1.25	1.67	1.71	0.17	0.25	0	0	0.52
1.1	500	0.01	0	0	0.15	0.68	0.95	1.15	1.06	0.13	0	0	0	0.42
1.2	500	0.04	0	0	0.21	0.62	0.45	1.04	1.25	0.16	0.04	0	0	0.29
1.3	500	0	0	0	0.18	0.29	0.29	0.59	1.01	0.12	0.01	0	0	0.30
1	1000	0.09	0	0	0.37	0.68	0.69	1.10	0.74	0.16	0.07	0	0	0.35
1.1	1000	0.04	0	0	0.22	0.42	0.42	0.41	0.49	0.09	0	0	0	0.22
1.2	1000	0	0	0	0.14	0.18	0.21	0.34	0.35	0.06	0	0	0	0.23
1.3	1000	0	0	0	0.12	0.06	0.12	0.15	0.24	0.06	0	0	0	0.18
10 experts														
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14
1	500	0.22	0	0	0.41	1.44	1.14	1.91	1.52	0.16	0.16	0	0	0.63
1.1	500	0.04	0	0	0.30	0.82	0.74	1.27	1.09	0.15	0.06	0	0	0.39
1.2	500	0.03	0	0	0.15	0.47	0.47	0.76	0.75	0.12	0	0	0	0.26
1.3	500	0	0	0	0.16	0.25	0.26	0.56	0.49	0.09	0	0	0	0.26
1	1000	0.13	0	0	0.18	0.72	0.68	0.88	0.91	0.16	0.07	0	0	0.27
1.1	1000	0.04	0	0	0.17	0.47	0.33	0.51	0.58	0.12	0.04	0	0	0.26
1.2	1000	0	0	0	0.15	0.11	0.19	0.31	0.33	0.10	0	0	0	0.16
1.3	1000	0	0	0	0.07	0.09	0.06	0.18	0.24	0.04	0	0	0	0.16

TAB. 6.13 – Pourcentages de violation du critère de risque 1% avec intégration d'un coefficient de sécurité au sein de l'approche SPSDR

L'introduction du coefficient de sécurité permet d'assurer plus facilement la satisfaction du critère 1%. Les violations sont moins nombreuses et moins importantes. Par contre, les bénéfices deviennent moins importants. Les bénéfices chutent logiquement aussi. Par exemple, avec 10 experts et une taille de 1000 scénarios par expert, le coût varie de 529 283 € sans coefficient de sécurité à 401 324 €, voire même 119 765 €. Cependant, il faut tout de même noter que ce bénéfice de 401 324 €, associé à un coefficient d'immunité de 1,1 est toujours plus important que le bénéfice des approches statiques sans réajustement empirique. L'utilisation d'un tel coefficient peut donc permettre de s'assurer un niveau de violation satisfaisant, tout en conservant un niveau de bénéfices assez intéressant.

6.7. RÉSULTATS NUMÉRIQUES AVEC SPSDR

1 expert																	
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
1	500			-685		051						-805		387			
1,1	500			-442		859						-616		584			
1,2	500			-382		579						-611		108			
1,3	500			-263		727						-513		198			
1	1000			-445		942						-649		509			
1,1	1000			-407		391						-588		435			
1,2	1000			-276		181						-388		829			
1,3	1000			-80		539						-378		721			
5 experts																	
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
1	500			-665		324						-829		663			
1,1	500			-502		418						-663		033			
1,2	500			-489		479						-600		676			
1,3	500			-340		927						-519		367			
1	1000			-526		201						-674		019			
1,1	1000			-366		736						-529		929			
1,2	1000			-242		318						-447		140			
1,3	1000			-102		007						-325		850			
10 experts																	
Immunité	Taille Experts	5	5	6	6	7	7	8	8	10	10	12	12	14	14	16	16
1	500			-714		686						-785		455			
1,1	500			-547		527						-689		168			
1,2	500			-428		546						-587		128			
1,3	500			-308		212						-499		954			
1	1000			-529		283						-642		826			
1,1	1000			-401		324						-573		866			
1,2	1000			-239		592						-465		459			
1,3	1000			-119		765						-360		135			

TAB. 6.14 – Coût moyen après intégration d'un coefficient de sécurité au sein de l'approche SPSDR

Remarque 6.3 Nous retrouvons les mêmes conclusions que [14] (voir 6.2.4). En effet, avec l'optimisation robuste, nous obtenons une bonne gestion des contraintes mais la performance économique peut être très faible. En revanche, SPSDR propose une approche qui, certes, est moins performante dans la gestion des contraintes (bien qu'elle obtienne des gestions satisfaisantes au prix de réglages de paramètres) mais qui permet de dégager des profits économiques intéressants. SPSDR propose donc un meilleur compromis risque-coûts que l'optimisation robuste.

6.7.4 Synthèse des résultats

Synthèse 6.4 (Synthèse Générale)

Satisfaction des objectifs

Avec un certain paramétrage, l'approche SPSDR permet de satisfaire le niveau maximal de violations de contrainte tout en maximisant les bénéfices ;

Violations des contraintes

1. *SPSDR n'a pas été conçue pour satisfaire des contraintes en probabilité : nous n'avons aucune assurance que le niveau de violations maximal soit satisfait ;*
2. *il existe des paramétrages qui permettent la satisfaction des contraintes mais le réglage nécessite de l'empirisme.*

Comportement du modèle

Pour réduire le nombre de violations en deçà du niveau maximal, il faut que les décisions soient le plus robustes possibles, c'est-à-dire qu'elles aient été construites à l'aide d'un maximum de scénarios possibles. Plusieurs possibilités pour ce faire :

1. *Avoir une taille d'expert suffisamment grande (mais pas trop pour que les temps d'optimisation demeurent raisonnables) ;*
2. *Réduire le nombre d'ensembles de décisions par jour. Cela permet aussi de réduire les temps importants de calculs provoqués par la partition des scénarios en ensembles de décisions ;*
3. *Accroître le nombre d'experts. Les temps de résolution sont augmentés par le nombre d'experts utilisé ;*
4. *Rajouter un coefficient de sécurité : il permet de forcer la minimisation des violations. Cependant, il faut veiller à ne pas trop dégrader le coût moyen.*

Coûts moyens en simulation

Sans rechercher de paramétrage optimal pour SPSDR, nous avons obtenu des gains moyens de 531 785 €, gains supérieurs aux approches statiques en utilisation normale (396 094 € pour le modèle robuste). D'autre part, ces gains sont de peu inférieurs aux gains obtenus en paramétrant optimalement les approches déterministe et robuste ($\approx 640\,000$ €). Nous dégagons ainsi un intérêt certain à utiliser cette approche par rapport aux approches statiques étudiées. Nous retrouvons les conclusions de l'article de référence [14] : SPSDR propose un meilleur compromis risque-coût que l'optimisation robuste.

Temps de calculs

L'essentiel des temps de calculs de SPSDR est concentré dans la création des partitions. Pour assurer un niveau de violation satisfaisant, il faut des experts de 1000 scénarios minimum. Le tableau 6.1 indique qu'il faut en moyenne 500s pour calculer une partition, soit $8 \times 2 \times 500 = 8000s$ en moyenne dans notre exemple, soit plus de 2H. Les temps de calculs de SPSDR sont donc relativement élevés. Cependant SPSDR peut être massivement parallélisée : les constructions de partition et les optimisations de chaque expert sont indépendants donc peuvent être traités en parallèle. Les gains de temps pourraient donc être réduits considérablement.

Bibliographie

- [1] R. Apparigliato, J. Thénier, and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : Application on a Hedging Problem in Electrical Industry. *11th International Conference on Stochastic Programming, Vienna (Austria)*, August 27-31 2007.
- [2] F. Babonneau, C. Beltran, O. du Merle, C. Tadjani, and J.-P. Vial. The Proximal Analytic Center Cutting Plane Method. Technical report, Logilab - HEC - University of Geneva, 2003.
- [3] N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to Use Expert Advice. *Journal of the ACM*, 44(3) :427–485, 1997.
- [4] O. du Merle and J.-P. Vial. Proximal ACCPM, a Cutting Plane Method for Column Generation and Lagrangian Relaxation : Application to the p-Median Problem. Technical report, Logilab - HEC - University of Geneva, 2002.
- [5] P. Hansen and N. Mladenovic. Variable Neighborhood Search. *Computers and Operations Research*, 24 :1097–1100, 1997.
- [6] P. Hansen, N. Mladenovic, and D. Perez-Brito. Variable Neighborhood Decomposition Search. *Journal of Heuristics*, 7 :335–350, 2001.
- [7] H. Heitsch and W. Römis. Scenario Reduction Algorithms in Stochastic Programming. *Computational optimization and applications*, 24 :187–206, 2003.
- [8] O. Kariv and L. Hakimi. An Algorithmic Approach to Network Location Problems. II : the p-Medians. *SIAM Journal of Applied Mathematics*, 37(3) :539–560, 1979.
- [9] Y. Nesterov and J.-Ph. Vial. Confidence Level Solutions for Stochastic Programming. *Automatica*, 2008 (to appear).
- [10] G. Ch. Pflug. Scenario Tree Generation for Multiperiod Financial Optimization by Optimal Discretization. *Mathematical programming*, 89, number 2 :251–271, 2001.
- [11] P. Raghavan and C.D. Thompson. Randomized Rounding : a Technique for Provably Good Algorithms and Algorithmic Proofs. *Combinatorica*, 7(4) :365–374, 1987.
- [12] J. Reese. Methods for Solving the p-Median Problem : An Annotated Bibliography. Technical report, Department of Mathematics, Trinity University, August 11, 2005.
- [13] A. Shapiro and A. Nemirovski. On Complexity of Stochastic Programming Problems. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*. Springer, London, 2005.
- [14] J. Thénier and J.-P. Vial. Step Decision Rules for Multistage Stochastic Programming : a Heuristic Approach. *Automatica*, 2008 (to appear).

BIBLIOGRAPHIE

Chapitre 7

Les processus stochastiques du problème de couverture

Sommaire

7.1	Introduction	208
7.2	La marge physique de production	209
7.2.1	Définition	209
7.2.2	Estimation de la marge avec des aléas gaussiens	210
7.2.3	Estimation de l'offre de production thermique	211
7.2.4	Estimation de l'offre de production hydraulique	214
7.2.5	Estimation probabiliste de la marge	220
7.2.6	Résultat typique	222
7.2.7	Estimation de modèles ARMA	223
7.3	Les prix de marché	224
7.3.1	Introduction	224
7.3.2	Définitions	225
7.3.3	Du modèle factoriel unitaire au modèle factoriel par produit	226
7.3.4	Discrétisation en temps du modèle à 2 facteurs "par produit"	230
7.3.5	Approximation sur les facteurs de forme	231
7.3.6	Création de scénarios	232
	Bibliographie	232

Ce chapitre regroupe les méthodes de génération des différents processus stochastiques nécessaires au problème de couverture contre le risque de défaillance physique. Le calcul de la marge est actuellement très simplifié, il suppose que tous les aléas sont gaussiens et que toutes les centrales peuvent produire à leur maximum. Disposant de générateurs d'aléas d'hydraulicité et d'indisponibilité, nous proposons une méthodologie d'intégration de ces aléas au sein du calcul de marge pour le rendre plus robuste. D'autre part nous améliorons le calcul de marge hydraulique en prenant en compte toutes les contraintes techniques de gestion d'une vallée. Ce calcul a été intégré à un outil opérationnel [21]. Une note interne [17] a été diffusée sur ce sujet. Un article a été soumis au journal "IEEE Transaction on Power Systems" [20]. Les scénarios de prix sont générés à partir d'un modèle à deux facteurs classique en finance mis à disposition par EDF. Nous décrirons brièvement son principe.

7.1 Introduction

Le chapitre 5 présente une façon d'optimiser les décisions de couverture, dans le but d'éviter le risque de défaillance physique. La modélisation proposée sépare la gestion de la production électrique de l'optimisation de la couverture contre le risque physique. Cependant, pour conserver une cohérence entre les deux phases d'optimisation, nous introduisons la notion de marge de production. De ce fait, la deuxième phase est modélisée comme un problème de minimisation du coût de la couverture sous la contrainte d'avoir une marge suffisante pour satisfaire un critère de risque particulier. Ce problème est un problème d'optimisation stochastique où les variables aléatoires sont la marge de production et les prix de marché.

La marge de production est le concept de base de la formulation du problème de couverture. Cette notion est déjà utilisée actuellement en exploitation mais calculée très simplement. Par définition, la marge est la différence entre l'offre totale et la demande. Jusqu'à présent, l'offre totale est considérée comme la somme des puissances maximales disponibles (aucune contrainte technique n'est prise en compte) alors que les aléas (demande, hydraulicité, indisponibilités fortuites) sont simplement supposés gaussiens. De plus ces calculs proviennent de modèles utilisés à l'horizon moyen-terme, sans considération avec les outils de l'horizon court-terme. Ce chapitre propose une nouvelle approche pour calculer la marge, notamment en intégrant des aléas plus réalistes (à l'aide de générateurs existants) dans le calcul de la capacité maximale de production. Quant à la capacité de production maximale hydraulique, la modélisation a été affinée en considérant notamment les contraintes d'influencement de chaque vallée. Désormais, il est ainsi nécessaire de résoudre le problème d'optimisation linéaire de gestion de vallée pour estimer sa capacité de production, selon un scénario d'hydraulicité. Pour la capacité thermique, un modèle exponentiel a été adapté pour modéliser l'aléa d'indisponibilité. Enfin, la demande est caractérisée à l'aide de scénarios de température fournis par Météo France. La procédure mise en place et présentée dans ce chapitre est donc une première approche, assez originale et spécifique aux problématiques court-terme.

L'autre processus stochastique, dont l'obtention de scénarios est décrite dans ce chapitre, est le processus de prix de marché. Pour cela, nous appliquons un modèle à 2 facteurs classique en finance. Un tel modèle, que nous décrirons rapidement, a déjà été implémenté à EDF [9] et a été mis à la disposition pour cette thèse par le groupe "Risque et Valorisation".

7.2 La marge physique de production

7.2.1 Définition

Soit un ensemble d'actifs physiques de production P constitué d'unités thermiques et hydrauliques, respectivement notées P^θ et P^η . Nous notons $K = J(24/\Delta t)$ le nombre de pas de temps de l'horizon de travail, constitué de J jours échantillonnés au pas de temps Δt , exprimé en heures (les valeurs typiques pour l'hebdomadaire sont : $J = 12$ jours, $\Delta t = 2$ heures, $K = 144$ pas de temps). Ce parc dispose d'une offre prévisionnelle totale en puissance, évaluée à l'instant t_0 pour tout instant $k > t_0$, $k = 1, \dots, K$, notée O_k et définie comme le cumul des offres des parcs P^θ et P^η , notés respectivement O_k^θ et O_k^η , soit $O_k = O_k^\theta + O_k^\eta$. Pour évaluer le risque de défaillance, nous définissons la marge prévisionnelle de production M_k , évaluée à l'instant t_0 pour tout instant $k > t_0$, $k = 1, \dots, K$, comme l'écart entre l'offre totale du parc de production O_k et la demande prévisionnelle totale de consommation électrique, notée D_k : $M_k = O_k - D_k$, $k = 1, \dots, K$, exprimée en MW. La marge de production dépendant des aléas affectant le système électrique, nous notons :

$$M_k(\xi_\theta, \xi_\eta, \xi_\delta) = O_k(\xi_\theta, \xi_\eta) - D_k(\xi_\delta) \quad (7.1a)$$

$$= O_k^\theta(\xi_\theta) + O_k^\eta(\xi_\eta) - D_k(\xi_\delta), \quad k = 1, \dots, K, \quad (7.1b)$$

pour indiquer qu'elle constitue un vecteur aléatoire dépendant de l'aléa de disponibilité du parc thermique, noté ξ_θ , qui va influencer notablement sur l'offre du parc thermique, de l'aléa sur les apports hydrauliques aux réservoirs, noté ξ_η , qui va jouer sur l'offre du parc hydraulique et de l'aléa sur la demande de consommation électrique, noté ξ_δ . Si $\Omega = \{\omega_1, \dots, \omega_n\}$ est l'espace de probabilité associé à M , supposé fini, nous notons $\{M \in \mathbb{R}^K\} = \{\omega \in \Omega : M(\omega) \in \mathbb{R}^K\}$ et de façon similaire pour les autres vecteurs aléatoires :

$$\begin{aligned} \{O^\theta \in \mathbb{R}_+^K\} &= \{\omega_\theta \in \Omega_\theta : O^\theta(\omega_\theta) \in \mathbb{R}_+^K\}, \\ \{O^\eta \in \mathbb{R}_+^K\} &= \{\omega_\eta \in \Omega_\eta : O^\eta(\omega_\eta) \in \mathbb{R}_+^K\}, \\ \{D^\delta \in \mathbb{R}_+^K\} &= \{\omega_\delta \in \Omega_\delta : D^\delta(\omega_\delta) \in \mathbb{R}_+^K\}. \end{aligned}$$

Nous proposons ici, comme le fait l'exploitant, de n'estimer la marge qu'aux heures de pointe. La marge de production est évaluée, pour chaque jour de la période d'étude, à l'heure de pointe du jour, définie comme l'instant de ce jour où la demande de consommation atteint sa valeur maximale. Pour déterminer les heures de pointe, nous considérons un scénario donné de demande de consommation $D \in \mathbb{R}_+^K$ que nous découpons en J périodes d'une journée :

$$D = \left[\{D\}_1^{(\frac{24}{\Delta t})} \dots \{D\}_{(j-1)(\frac{24}{\Delta t})}^{j(\frac{24}{\Delta t})} \dots \{D\}_{(J-1)(\frac{24}{\Delta t})}^{J(\frac{24}{\Delta t})} \right], \quad \forall j = 1, \dots, J. \quad (7.2)$$

Nous notons pour simplifier $D_j = \{D\}_{(j-1)(\frac{24}{\Delta t})}^{j(\frac{24}{\Delta t})}$, $j = 1, \dots, J$. Pour chaque jour de la période d'étude, nous déterminons la pointe de la demande de consommation, ce qui donne un vecteur

$\tilde{D} \in \mathbb{R}_+^{K/J}$ défini par $\hat{D} = [\|D_1\|_\infty \dots \|D_j\|_\infty \dots \|D_J\|_\infty]$ et à partir duquel nous déduisons les J heures de pointe. Ainsi, avec les 51 scénarios de demande de consommation électrique, nous obtiendrions 51 scénarios d'heures de pointe. Dans la pratique, nous admettons que la détermination des heures de pointe est déterministe et réalisée à partir du scénario moyen de demande. Dans ces conditions, la marge de production à l'heure de pointe, prévisionnelle pour les J prochains jours est définie par :

$$M_j(\xi_\theta, \xi_\eta, \xi_\delta) = O_j(\xi_\theta, \xi_\eta) - D_j(\xi_\delta), \quad j = 1, \dots, J. \quad (7.3)$$

Ainsi, le calcul de la marge nécessite d'estimer l'offre du parc de production thermique et celle du parc de production hydraulique.

7.2.2 Estimation de la marge avec des aléas gaussiens

Actuellement, le calcul de la marge de production est effectuée en supposant que toutes les centrales sont disponibles à leur capacité maximale à tout instant et que les aléas sont gaussiens. Nous pouvons écrire une version simple de ce calcul.

La capacité maximale de production thermique au jour j est la somme des capacités maximales disponibles en ce même jour. Nous notons $\alpha_{i_\theta}^j(\xi_\theta)$ la variable aléatoire représentant le coefficient de disponibilité de la centrale i_θ au jour j et dépendant de l'aléa d'indisponibilité ξ_θ . L'offre du parc thermique est donc :

$$O_j^\theta = \sum_{i_\theta=1}^{N_\theta^j} \alpha_{i_\theta}^j(\xi_\theta) \bar{P}_{i_\theta}^j, \quad \forall j = 1, \dots, J,$$

avec N_θ^j le nombre de centrales thermiques disponibles au jour j et $\bar{P}_{i_\theta}^j$ la puissance maximale de la centrale i_θ au jour j .

Pour la capacité maximale de production hydraulique au jour j , nous agrégeons toutes les réserves en un "lac France". L'offre du parc hydraulique est donc :

$$O_j^\eta = \rho_j V_j(\xi_\eta), \quad \forall j = 1, \dots, J,$$

avec ρ_j le rendement moyen du parc hydraulique et $V_j(\xi_\eta)$ la quantité d'eau maximale disponible au jour j , soumis à l'aléa d'hydraulicité ξ_η .

Soit $D_j(\xi_\delta)$ la demande et ξ_δ l'aléa sur cette demande.

La marge de production s'écrit donc :

$$M_j(\xi_\theta, \xi_\eta, \xi_\delta) = \sum_{i_\theta=1}^{N_\theta^j} \alpha_{i_\theta}^j(\xi_\theta) \bar{P}_{i_\theta}^j + \rho_j V_j(\xi_\eta) - D_j(\xi_\delta), \quad \forall j = 1, \dots, J.$$

La réglementation en vigueur impose une couverture contre le risque de défaillance physique à $1-p$, ce qui peut se formuler sous la forme d'une contrainte en probabilité :

$$Prob(\sum_{i_\theta=1}^{N_\theta^j} a_{i_\theta}^j(\xi_\theta) \bar{P}_{i_\theta}^j + \rho_j V_j(\xi_\eta) - D_j(\xi_\delta) \geq 0) \geq 1 - p, \forall j = 1, \dots, J. \quad (7.4)$$

Comme les aléas sont supposés gaussiens, chaque $a_{i_\theta}^j$ suit une loi normale $\mathcal{N}(\mu_{j,i_\theta}, \sigma_{j,i_\theta}^2)$, V_j suit une loi normale $\mathcal{N}(\mu_{j,\eta}, \sigma_{j,\eta}^2)$ et D suit une loi normale $\mathcal{N}(\mu_{j,\delta}, \sigma_{j,\delta}^2)$.

A l'aide de [10], nous obtenons l'équivalent déterministe de la contrainte en probabilité (7.4) :

$$\sum_{i_\theta=1}^{N_\theta^j} \mu_{j,i_\theta} \bar{P}_{i_\theta}^j + \rho_j \mu_{j,\eta} - \mu_{j,\delta} + \Phi^{-1}(1-p) \sqrt{\sum_{i_\theta=1}^{N_\theta^j} \sigma_{j,i_\theta}^2 (\bar{P}_{i_\theta}^j)^2 + \rho_j^2 \sigma_{j,\eta}^2 + \sigma_{j,\delta}^2} \geq 0, \quad \forall j = 1, \dots, J,$$

où Φ est la fonction de répartition de la loi gaussienne.

Nous venons donc de définir la marge sous l'hypothèse d'aléas gaussiens.

7.2.3 Estimation de l'offre de production thermique

En l'absence d'aléa

Le parc thermique P^θ est composé de près de 58 tranches nucléaires et de 50 unités thermiques classiques (charbon, houille, fioul et turbines à gaz). De façon simplifiée, nous admettons que l'offre O^θ de production thermique est la capacité maximale de production de P^θ c'est-à-dire la puissance installée totale :

$$O_j^\theta = \sum_{i \in \text{card}(P^\theta)} \bar{P}_j^i, \quad (7.5)$$

où \bar{P}_j^i est la puissance installée de l'unité i disponible au temps j . En réalité, pour évaluer précisément la capacité maximale de production d'une centrale thermique, il faudrait tenir compte de son état et de ses contraintes de fonctionnement. Dans la pratique, un tel calcul est d'une complexité similaire au problème du placement de la production thermique résolu par programmation dynamique. C'est pourquoi, nous admettons qu'en prenant les dispositions nécessaires, une centrale thermique peut produire au maximum de sa puissance installée, sans tenir compte de ses contraintes, à une heure définie suffisamment à l'avance. Ainsi, en l'absence d'aléa, nous supposons que l'offre O^θ de production thermique du parc thermique P^θ , constituée de $I = \text{card}(P^\theta)$ unités, peut être estimée par la relation (7.5).

Modélisation de l'aléa d'indisponibilité fortuite du parc thermique

Dans la réalité, l'offre du parc thermique est soumise à l'aléa d'indisponibilité fortuite affectant les centrales. Pour en tenir compte, nous allons introduire un modèle simplifié d'indisponibilité fortuite fondé sur une loi exponentielle [1, 2, 16].

Fondements Dans le cas basique d'une panne de durée ($\tau = 0$), nous supposons que la probabilité que la centrale soit en état de marche pendant une durée d supérieure à un temps t est $\mathbb{P}\{d > t\} = e^{-\alpha t}$, $t > 0$. Le paramètre α représente le taux de panne. Nous avons $\mathbb{P}\{d \leq t\} = 1 - e^{-\alpha t}$ et la densité de probabilité associée est $f(t) = \alpha e^{-\alpha t}$. Une propriété fondamentale de la loi exponentielle est qu'elle se réinitialise automatiquement. Si nous savons que jusqu'à l'instant τ il n'y a pas eu de panne, la probabilité que la durée de fonctionnement dépasse τ est $e^{-\alpha(t-\tau)}$: tout se passe comme si la loi se remettait à zéro à l'instant τ . Dans le cas discret, nous avons $\mathbb{P}\{d \geq n\} = e^{-\alpha n}$ avec n la durée de fonctionnement sans panne, exprimée en pas de temps.

Pour une panne durant un temps τ non nul, si la panne se produit à l'instant t , elle se finit à l'instant $t + \tau$ et la durée de fonctionnement de la centrale suit de nouveau une loi exponentielle réinitialisée à l'instant $t + \tau$. Dans le cas discret, si la panne se produit à l'instant n et que la réparation commence à l'instant $n + 1$ pour une durée de m pas de temps, la centrale est indisponible pendant la période $[n + 1, \dots, n + m]$. A l'instant $n + m + 1$, la centrale est de nouveau en activité, mais elle peut tomber en panne le jour même de son retour. Nous réinitialisons la loi exponentielle avec $n = 1$: $\mathbb{P}\{d \geq 1\} = e^{-\alpha}$. La centrale peut donc tomber en panne dès l'instant de son retour avec la probabilité $\mathbb{P}\{d < 1\} = 1 - e^{-\alpha}$.

Remarque 7.1 *Il est possible de calculer les probabilités de panne. Les calculs sont détaillés en Annexe C.*

Estimation du taux de panne La détermination du paramètre α de la loi exponentielle, en fonction du nombre moyen de pannes observées par an, peut se faire selon deux procédés : un procédé plus précis que le second mais nécessitant davantage de calculs que le procédé simplifié mais suffisamment précis en pratique.

- **Calcul précis.** Nous notons X la variable aléatoire qui indique le nombre de pannes sur une année. La loi de X est connue et T le nombre de pas de temps de 2h dans une année ($T = 365 \times 12 = 4380$). Nous savons déterminer la probabilité de n'avoir aucune panne en T pas de temps, d'en avoir une seule panne, d'en avoir deux pannes, etc. L'espérance de X est alors connue en fonction de α :

$$\mathbb{E}(X) = \sum_{k \geq 0} k \mathbb{P}\{X = k\}.$$

Comme nous avons déterminé qu'il y a en moyenne n pannes par an, nous résolvons l'équation en α : $\mathbb{E}(X) = n$.

- **Calcul simplifié.** La méthode de calcul simplifiée consiste à dire que s'il y a n pannes par an, cela correspond à T/N pas de temps sans panne en moyenne, T étant le nombre de pas de temps de 2 heures dans une année. Comme la loi exponentielle a pour fonction de répartition $F(t) = 1 - e^{-\alpha t}$, elle a pour densité $f(t) = \alpha e^{-\alpha t}$ et pour espérance :

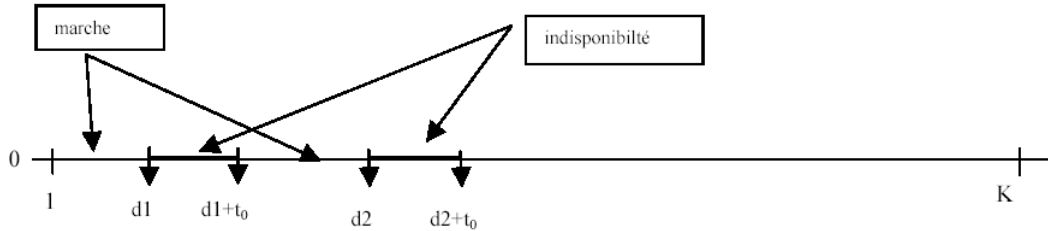
$$\mathbb{E}_\alpha = \int_0^{+\infty} t f(\alpha, t) dt = \frac{1}{\alpha}.$$

Le taux de panne α peut donc être évalué par : $\alpha = \frac{n}{T}$.

Génération de M_θ scénarios de disponibilité thermique

Pour tenir compte de cet aléa de pannes fortuites des unités thermiques, nous allons calculer un grand nombre de fois, disons M_θ fois, cette capacité de production thermique, en générant M_θ scénarios de disponibilité thermique à l'aide du modèle d'indisponibilité fortuite présenté précédemment.

Nous travaillons sur un horizon avec K pas de temps, échantillonné à Δt . A l'instant 0, la centrale fonctionne. Nous tirons un premier nombre selon la loi exponentielle. Il va nous indiquer l'instant de la première panne (nous limitons la loi à l'intervalle $1, \dots, K$). Nous ajoutons la durée de la panne ; si la date obtenue est inférieure à K , nous recommençons avec une loi exponentielle réinitialisée.



Les étapes du processus de tirage sont les suivantes :

- d_1 = tirage aléatoire ;
- $d_1 + t_0$ = fin de la panne ;
- d_2 = second tirage aléatoire $> d_1 + t_0$;
- $d_2 + t_0$ = fin de la panne, etc.

La centrale est en activité sur les intervalles : $[1, d_1]$, $[d_1 + t_0, d_2]$, $[d_2 + t_0, d_3]$ et ainsi de suite.

Nous disposons ainsi d'une matrice $\hat{C}_{\max}^T = \left[\left(\hat{C}_{\max}^{\theta} \right)_{s,j} \right]$, $\hat{C}_{\max}^T \in \mathbb{R}_+^{M_{\theta} \times J}$, où chaque vecteur-colonne est considéré comme les M_{θ} réalisations d'un processus aléatoire. Notons que le modèle d'indisponibilité fortuite prend en compte les contraintes initiales sur le parc thermique.

7.2.4 Estimation de l'offre de production hydraulique

Le parc hydraulique est organisé en une quarantaine de vallées nationales, chaque vallée étant modélisée comme un ensemble cohérent de réserves hydrauliques et d'usines de production, interconnectées. Il représente plus de 100 usines hydrauliques. Chaque vallée est caractérisée par sa topologie (cf. figure 7.1) qui décrit les interconnexions entre réserves et usines, leurs nature (usine de pompage ou de turbinage) et certains paramètres comme le délai de parcours de l'eau entre deux réservoirs. Pour chaque usine et pour chaque réserve hydraulique, sont donnés :

- un ensemble de contraintes à respecter telles que les volumes d'eau minimaux et maximaux des réservoirs, les limitations de turbinés et/ou de puissance ;
- une structure de coût définie à partir d'une valeur d'usage de l'eau, calculée par des techniques de gestion de stock à l'horizon annuel et permettant de donner une vision du futur. La valeur d'usage de l'eau permet de définir le coût de production hydraulique associé à une vallée V ;
- un ensemble de courbes débit/puissance permettant de déterminer la puissance obtenue pour un débit donné, compte tenu du rendement de l'usine et de son niveau de production.

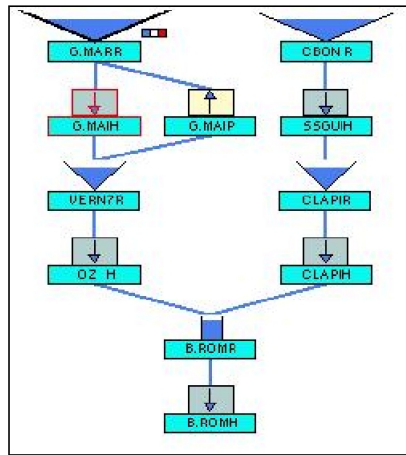


FIG. 7.1 – Exemple de représentation d'une vallée (ici la Romanche)

En l'absence d'aléas

Dans un premier temps, nous allons poser le problème de l'estimation de la capacité maximale de production d'une vallée, puis celui d'un ensemble de vallées. Nous décrirons ensuite la procédure retenue qui aboutit à une évaluation pessimiste de l'offre hydraulique mais aboutissant à des

temps de calcul admissibles (une évaluation plus réaliste de l'offre hydraulique par une technique de tirage de type Monte-Carlo, conduirait à des temps de calculs rédhibitoires).

Position du problème

– Cas d'une vallée.

L'estimation de la capacité maximale de production d'une vallée est un problème difficile. La raison provient du fait qu'il faut tenir compte à la fois de la topologie de la vallée et de ses contraintes (influencement, cotes d'exploitation des réservoirs, délai de parcours de l'eau, ...) et de l'effet-mémoire de la vallée : à un instant donné, la capacité maximale de production d'une vallée dépend de la façon dont cette vallée a été gérée avant. Il existe donc une grande combinatoire. Nous allons donc procéder en deux étapes : nous déterminons le problème du calcul de la production d'une vallée compte tenu de sa topologie et de ses contraintes puis nous chercherons à maximiser cette production pour évaluer sa capacité maximale de production.

Etape 1 : Production coordonnée d'une vallée.

Nous nous plaçons dans un contexte de résolution du problème du placement à coût optimal de la production thermique et hydraulique, c'est-à-dire de l'élaboration des plannings optimaux de production thermique et hydraulique satisfaisant une demande de consommation et respectant les contraintes de fonctionnement propres aux différentes unités de production. Nous sommes obligés de considérer le parc dans son ensemble car le nombre d'unités de production, leurs structures de coût, leur fonctionnement (la production selon certains paliers) ainsi que la nature couplante de la contrainte de satisfaction de la demande rendent le fonctionnement de chaque unité de production interdépendant les uns des autres. Ces contraintes confèrent à ce problème d'optimisation des caractéristiques difficiles : grande taille, fonction non linéaire, non convexe, non différentiable. En se fondant sur la théorie de la dualité et en appliquant une technique de décomposition par les prix [5], [11], possible en raison de la structure additive du problème, le problème primal initial de minimisation d'un coût est transformé par relaxation lagrangienne en un problème dual de maximisation d'une fonction duale [7], [3]. Les contraintes couplantes de satisfaction de la demande en puissance sont traitées par pénalisation. Tout écart entre demande et production est facturé à un certain coût. Des courbes de pénalités sont élaborées pour introduire un coût de défaillance (respectivement de sur-production) d'autant plus élevé que le niveau de défaillance (respectivement de sur-production) est important. Le problème se rapporte donc à un ensemble de sous-problèmes thermiques et hydrauliques plus petits, résolus selon des méthodes adaptées à chaque sous-problème, le tout étant supervisé par un algorithme de coordination de la production. Cet algorithme de coordination a pour fonctions de déterminer une chronique de prix transmise aux unités afin de les inciter à produire (rémunération de la production), de calculer des sous-gradients pour approcher la fonction duale, d'ajuster la chronique de prix afin de respecter la demande et de maximiser finalement la fonction duale approchée. La chronique de prix transmise à chaque unité l'incite à une certaine production en fonction de ses caractéristiques et de ses contraintes. Les sous-problèmes sont résolus, pour chaque unité, par une technique de programmation dynamique concernant le parc thermique et, pour les vallées hydrauliques, par une technique de programmation linéaire en phase de lagrangien simple et quadratique en phase de lagrangien augmenté (algorithme de points intérieurs primal-dual). Toute cette méthodologie permet à un modèle, appelé Apogée, de résoudre le problème global de gestion

de production avec une vision déterministe des aléas.

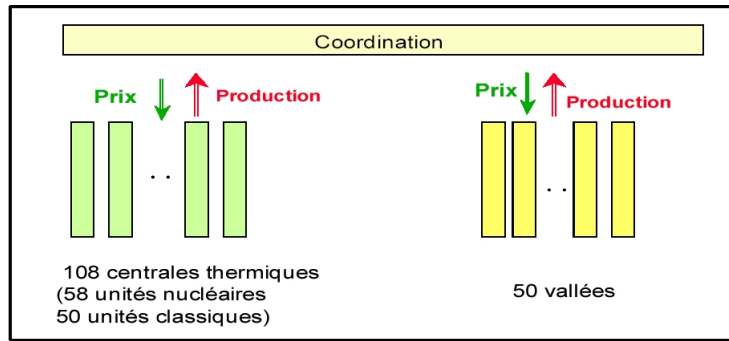


FIG. 7.2 – Contexte de résolution : décomposition par les prix.

Dans ce contexte, en ne tenant compte que des contraintes d'évolution des réservoirs (équations de flot), des contraintes de bornes sur les turbinés de chaque groupe et des contraintes de bornes sur les volumes des réservoirs, le problème du placement optimal sur critère économique de la production hydraulique d'une vallée, en phase de lagrangien simple, et pour une itération de l'algorithme général de coordination thermique et hydraulique par la méthode des faisceaux, est un programme linéaire $\min_x c^t x$, $Ax = b$, $x \geq 0$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ [18], [19]. Le problème hydraulique à résoudre a déjà été présenté au cours de cette thèse au chapitre 4. Dans cette formulation, conforme à celle du modèle opérationnel Apogée, nous ne tenons pas compte du caractère discret de la production hydraulique et de certaines contraintes comme les contraintes de non-changement de sens ou de gradient de débit qui confèreraient au problème une nature combinatoire et dont la résolution, à l'horizon journalier fait l'objet d'investigations [14]. De plus, seul le placement en puissance est traité puisque les réserves (obligation réglementaire de conserver de la puissance disponible, par exemple à produisant sur certaines centrales moins que la puissance maximale) ne sont pas prises en compte.

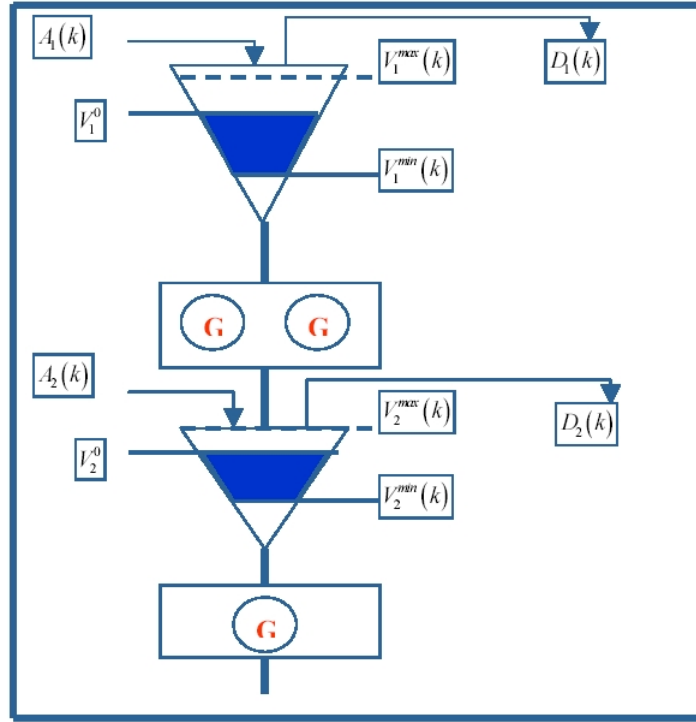


FIG. 7.3 – Modélisation d'une vallée et paramètres d'exploitation des réservoirs.

Etape 2 : Production maximale d'une vallée.

Une vallée hydraulique peut donc être considérée comme un système de production réagissant à un signal de prix $\lambda \in \mathbb{R}^K$. Nous allons donc, par un signal de prix approprié, inciter la vallée à produire au maximum de sa capacité, compte tenu de son état et de ses contraintes. Pour cela, nous effectuons préalablement un placement coordonné des moyens thermiques et hydrauliques à l'aide du modèle Apogée pour la demande de consommation électrique moyenne. Après convergence, nous récupérons la chronique des coûts marginaux du système (soit, en fait, la chronique des prix à la convergence ou, vu autrement, la solution duale du problème où les contraintes couplantes sont dualisées). Soit $\lambda^* \in \mathbb{R}^K$ cette chronique. Nous allons nous servir de cette chronique de base pour construire un signal de rémunération de la vallée suffisamment attractif pour l'inciter à produire au maximum de ses capacités aux heures de pointe.

Pour inciter la vallée à produire au maximum à la première heure de pointe $H1(j = 1)$, nous construisons le signal de prix $\lambda^1 = \lambda^*$ puis, au pas de temps correspondant à la première heure de pointe, nous multiplions la valeur originelle par un grand nombre γ , disons $\gamma = 100$ ou 1000 . Nous avons ainsi :

$$\lambda^1 = [\{\lambda^*\}_{k < H1} \dots \{\gamma\lambda^*\}_{H1} \dots \{\lambda^*\}_{k > H1}], \quad \forall k = 1, \dots, K. \quad (7.6)$$

La rémunération étant très forte à la première heure de pointe, la vallée va normalement produire au maximum à cette heure. Pour ce faire, nous résolvons : $\min_x c_1^t(\lambda^1)x, Ax = b, x \geq 0$. La solution de ce programme linéaire est une chronique de production. Nous nous intéressons uniquement à la valeur maximale de cette chronique, obtenue à l'heure de pointe $H1$. Soit

$C_0^\eta(j=1)$ cette solution. Nous avons ainsi déterminé très facilement, la capacité maximale de production de la vallée pour la première heure de pointe.

Les choses se compliquent pour évaluer la capacité de production maximale pour l'heure de pointe du second jour $C_0^\eta(j=2)$ car tout dépend de l'état de la vallée antérieurement à l'évaluation. Il est clair que $C_0^\eta(j=2)$ dépend de l'état du système à l'instant $j=1$ et est conditionné par la production effective du premier jour, que nous notons $\tilde{C}_0^\eta(j=1)$. Notons cette dépendance $C_0^\eta[j=2|\tilde{C}_0^\eta(j=1)]$. Nous avons nécessairement $0 \leq \tilde{C}_0^\eta(j=1) \leq C_0^\eta(j=1)$. Puisqu'à l'instant t_0 , nous ne savons ce que nous produirons à $j=1$, nous allons discrétiser l'intervalle $[0, C_0^\eta(j=1)]$ en P niveaux, disons $P=5$ pour fixer les idées. Cette discrétisation correspond en quelque sorte à P hypothèses de production pour l'instant $j=1$. Le problème de cette discrétisation est qu'elle n'assure pas que les P capacités soient réalisables en pratique. Nous sommes donc contraints à discrétiser, non pas les capacités de production, mais les prix. Pour chacune des P hypothèses de prix, nous pouvons ainsi, au prix de P optimisations, obtenir les P hypothèses de capacité de production souhaitées.

Pour la seconde heure de pointe, il y a donc P états initiaux possibles d'offre hydraulique. A $j=2$, nous générons le signal de prix attractif $\lambda_{j=2}$; nous résolvons alors P^2 fois le programme linéaire réduit, à l'aide de P hypothèses $\lambda_{k,j=2}$, $k=1, \dots, P$ et des P hypothèses $\lambda_{n,j=1}$, $n=1, \dots, P$:

$$\min_{x_2 \geq 0} c_2^t(\lambda_{k,j=2})x_2, \quad A_2x_2 = b - A_1x_1[\lambda_{n,j=1}]. \quad (7.7)$$

Ce processus est répété pour les jours suivants. Nous résolvons donc P^J programmes linéaires qui deviennent de plus en plus petits en taille. L'effort de calcul reste donc modeste pour P raisonnable.

– Cas de plusieurs vallées.

Pour V vallées, le problème de l'estimation de l'offre hydraulique globale devient délicat. En effet, puisque nous avons P^J capacités de production pour une vallée, nous aurons P^{JV} capacités pour un parc hydraulique constitué de V vallées. Même en prenant trois scénarios de production à chaque pas de temps, vingt vallées et sept pas de temps, le nombre de possibilités atteint 6.10^{66} .

Nous pouvons illustrer cette approche sur l'exemple d'une vallée réduite de 3 pas de temps, 2 réservoirs et 3 groupes de production. Les données de ce problème sont les suivantes (volumes, apports et turbinés en m^3 à chaque pas de temps, la valeur de l'eau ω en euros par m^3 , rendement ρ en MWs/m^3) :

Réservoir	V_0	V_{min}	V_{max}	A	ω
1	120	0	400	15	25
2	130	0	400	12	20

TAB. 7.1 – Exemple : données des réservoirs

7.2. LA MARGE PHYSIQUE DE PRODUCTION

Groupe	ρ	T^{min}	T^{max}
1	0.64	0	55
2	0.74	0	65
3	0.69	0	80

TAB. 7.2 – Exemple : données des groupes hydrauliques

La valeur maximale du coût est 30 euros/MWh. Nous discrétisons ce prix en 5 pas réguliers. Après application de notre technique, nous avons obtenu seulement 39 possibilités différentes¹ les unes des autres. Nous nous rendons ainsi compte qu'à partir d'un certain coût, les contraintes sont saturées et il est impossible de produire plus. Même si le nombre de calculs est réduit, il demeure important, montrant ainsi qu'une telle approche semble difficile à mettre en œuvre en exploitation.

Estimation par une procédure de Monte-Carlo Une façon moins coûteuse d'estimer la capacité maximale de production d'une vallée consisterait à générer un grand nombre de scénarios de rémunération. Nous disposerions ainsi d'une distribution de valeurs de production et nous pourrions prendre l'espérance comme valeur représentative de l'offre de la vallée.

Procédure retenue Comme il est difficilement envisageable d'évaluer ainsi la marge hydraulique, nous proposons une procédure fondée sur une situation moyenne. Elle est critiquable mais présente l'avantage d'estimer raisonnablement la marge hydraulique. Cette procédure est la suivante. Nous nous plaçons toujours dans le contexte de décomposition par les prix. Nous avons effectué préalablement un placement coordonné des moyens thermiques et hydrauliques à l'aide du modèle Apogée pour la demande de consommation électrique moyenne et nous disposons de la chronique des coûts marginaux du système, $\lambda^* \in \mathbb{R}^K$. La construction du signal de rémunération de la vallée est effectuée pour que, vue de t_0 , la vallée soit incitée à produire au maximum à toute les heures de pointe $\{H_j\}$, $j = 1, \dots, J$. Nous ne privilégions donc aucune heure, aucune situation. Nous supposons ainsi, qu'en moyenne, nous avons besoin de produire au maximum à chaque heure de pointe de chaque jour de la période d'étude. Il s'agit donc d'une situation moyenne pour laquelle l'effort de production est réparti. Nous disposerons donc d'une minoration de la capacité réelle de production. Le signal de prix sera d'abord initialisé par $\lambda = \lambda^*$ puis, aux pas de temps correspondant à toutes les heures de pointe, nous multiplierons la valeur originelle en ces points par un grand nombre γ , $\gamma = 100$ ou 1000 . Nous aurons ainsi :

$$\lambda^j = [\{\lambda^*\}_{k_j < H_j} \dots \{\gamma \lambda^*\}_{H_j} \dots \{\lambda^*\}_{k_j > H_j}], \quad k_j = (j-1) \left(\frac{24}{\Delta t} \right), \dots, j \left(\frac{24}{\Delta t} \right), \quad j = 1, \dots, J. \quad (7.8)$$

¹nous pouvons les détailler : $\{0, 0, 0\}, \{0, 0, 41\}, \{0, 0, 76\}, \{0, 0, 114\}, \{0, 48, 0\}, \{0, 48, 41\}, \{0, 48, 50\}, \{0, 48, 89\}, \{0, 83, 0\}, \{0, 83, 11\}, \{0, 83, 49\}, \{0, 121, 0\}, \{0, 121, 11\}, \{0, 121, 49\}, \{48, 0, 0\}, \{48, 0, 41\}, \{48, 0, 50\}, \{48, 0, 89\}, \{48, 48, 0\}, \{48, 48, 4\}, \{48, 48, 42\}, \{48, 51, 0\}, \{48, 51, 38\}, \{48, 90, 0\}, \{48, 90, 38\}, \{83, 0, 0\}, \{83, 0, 11\}, \{83, 0, 49\}, \{83, 11, 0\}, \{83, 11, 38\}, \{83, 49, 0\}, \{83, 49, 38\}, \{139, 0, 0\}, \{139, 0, 11\}, \{139, 0, 49\}, \{139, 11, 0\}, \{139, 11, 38\}, \{139, 49, 0\}, \{139, 49, 5\}.$

Modélisation de l'incertitude sur les apports hydrauliques

Pour tenir compte des incertitudes d'estimation sur les apports hydrauliques, nous allons supposer que toute valeur d'apport hydraulique est une variable aléatoire gaussienne de moyenne μ et d'écart-type σ (ce choix est validé par des études internes). Nous souhaitons générer des scénarios d'apports hydrauliques selon une loi normale bi-tronquée pour limiter les valeurs admissibles des apports à un intervalle de variation de bornes α et β . Nous cherchons à déterminer les deux paramètres de cette loi, la moyenne m et l'écart-type s , en fonction des paramètres μ et σ de la loi initiale.

La moyenne m de la loi bi-tronquée est :

$$m = \mu + \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \left[e^{-\frac{1}{2}\left(\frac{\alpha-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2} \right].$$

La variance se calcule par la formule $s^2 = \mathbb{E}(u^2) - m^2$, d'où :

$$s^2 = \sigma^2 + m(\mu - m) + \frac{1}{F_{\alpha\beta}} \frac{\sigma}{\sqrt{2\pi}} \left[\alpha e^{-\frac{1}{2}\left(\frac{\alpha-\mu}{\sigma}\right)^2} - \beta e^{-\frac{1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2} \right].$$

Les détails de ces calculs se trouvent dans l'annexe C.

Génération de N_η scénarios d'apports hydrauliques

Pour tenir compte de l'incertitude sur les apports hydrauliques, nous perturbons les valeurs nominales d'apports selon une loi gaussienne bi-tronquée. Nous disposons, comme pour la marge thermique, d'une matrice $\hat{C}_{\max}^H = \left[\left(\hat{C}_{\max}^\zeta \right)_{s,j} \right]$, $\hat{C}_{\max}^H \in \mathbb{R}_+^{N_\zeta \times J}$, où chaque vecteur-colonne est considéré comme les M_η réalisations d'un processus aléatoire. Dans la pratique, pour 7500 scénarios Apogée, le temps de calcul pour la marge hydraulique est d'environ 17 heures sur une machine Linux. On peut toutefois remarquer qu'en parallélisant le code, les temps de calcul pourraient être considérablement réduits.

7.2.5 Estimation probabiliste de la marge

Tirage de Monte Carlo

Nous disposons de M_δ scénarios de demande de consommation électrique (dans la pratique 51 scénarios), M_θ scénarios de disponibilité des centrales thermiques et M_η scénarios d'apports hydrauliques. Il existe donc $M_\delta M_\theta M_\eta$ estimations possibles de la marge de production. Comme ce nombre est très élevé, nous nous contentons d'évaluer la marge qu'en un nombre $n \ll M_\delta M_\theta M_\eta$ de scénarios ; il est donc nécessaire d'utiliser des techniques de tirage de scénarios afin d'obtenir un échantillon représentatif. De manière à parcourir au mieux l'espace des configurations avec un

nombre donné de simulations, la méthode de l'hypercube latin de McKay, Conover et Beckman (1979) est intéressante. Cette méthode consiste à diviser le domaine de variation de chaque variable d'entrée indépendamment en un nombre d'intervalles de même probabilité, puis à choisir au hasard une valeur dans chacun de ces intervalles. Le nombre d'intervalles est fixé au départ et identique pour chaque variable ; il correspond donc à la taille de l'échantillon. L'association des valeurs correspondant à chaque paramètre, et définissant une exécution du modèle numérique, est alors effectuée de manière aléatoire. Cette façon de procéder permet de parcourir l'espace multidimensionnel sans redondance. Cela ne signifie pas néanmoins que l'intégralité du domaine est couverte. Une seconde façon de procéder est d'effectuer le tirage aléatoire (technique de Monte-Carlo). En combinant les différents scénarios de base entre eux selon cette technique, nous disposons de n calculs de marge prévisionnelle vue de t_0 , notés (m_1, \dots, m_n) .

Estimation de la densité de probabilité de la marge

Nous supposons que la loi de probabilité de la marge prévisionnelle de production, variable aléatoire réelle, admet une densité par rapport à la mesure de Lebesgue λ sur \mathbb{R} , dite densité de probabilité, notée f . Nous cherchons à estimer cette densité à partir des n valeurs de marge prévisionnelle de production (m_1, \dots, m_n) . Considérer que (m_1, \dots, m_n) est un échantillon observé d'une loi à densité inconnue f suppose que ces observations sont les valeurs prises simultanément par n variables aléatoires (M_1, \dots, M_n) indépendantes et suivant cette loi de densité f . Trois techniques sont envisageables pour l'estimation, délicate, de la densité inconnue f :

- La première technique possible d'estimation est l'histogramme. Malheureusement, à taille d'échantillon donné, nous ne pouvons avoir qu'une idée assez vague de f . En effet, s'il existe des règles simples pour la détermination du nombre de classes, il est délicat de déduire la nature de la loi de la forme de l'histogramme ;
- La seconde technique est l'estimation paramétrique. Nous supposons que l'échantillon considéré suit une loi à densité inconnue mais qui est une loi normale. Comme les lois normales sont entièrement déterminées par la connaissance de deux paramètres, il suffit d'estimer ces paramètres. Malheureusement, à taille d'échantillon donné, nous ne pouvons faire d'estimation exacte (les résultats exacts sont nécessairement asymptotiques) ;
- La troisième technique est l'estimation non paramétrique, dite méthode à noyau, actuellement la plus efficace et la plus fréquemment utilisée pour estimer une densité [15, 6, 8]. L'estimation de la fonction de densité continue f , à partir de l'échantillon discret (m_1, \dots, m_n) , supposé issu de cette loi est réalisée sans faire d'hypothèses a priori sur l'appartenance de f à une famille de lois connues. L'estimation ne concerne donc plus un paramètre dans cette famille de loi, mais directement la fonction elle-même, d'où le terme de non paramétrique. Cette technique peut être considérée comme une amélioration significative de la technique de l'histogramme par introduction d'une fenêtre h et d'un noyau K , en général continu, et ayant de "bonnes propriétés" de régularité et de symétrie, permettant en final de disposer d'une bonne estimation de f (cf. Annexe D).

Finalement, nous avons choisi d'évaluer la densité par la méthode à noyau et fenêtrage adapté selon la taille de l'échantillon. L'estimation s'écrit :

$$\hat{f}_h(m, K) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{m - m_i}{h}\right).$$

Il est largement admis (et partiellement démontré) que le choix de h a plus d'influence que le choix de K . L'algorithme de mise en oeuvre est celui proposée par [4] avec noyau gaussien défini par :

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}, \quad -\infty < u < \infty.$$

7.2.6 Résultat typique

À titre d'illustration, un résultat typique d'estimation de la marge, obtenu par la démarche proposée avec des données réelles, est consigné à la figure 7.4 (données relatives à la semaine du 23 mars 2006). Les résultats complets sont consignés à l'annexe E.

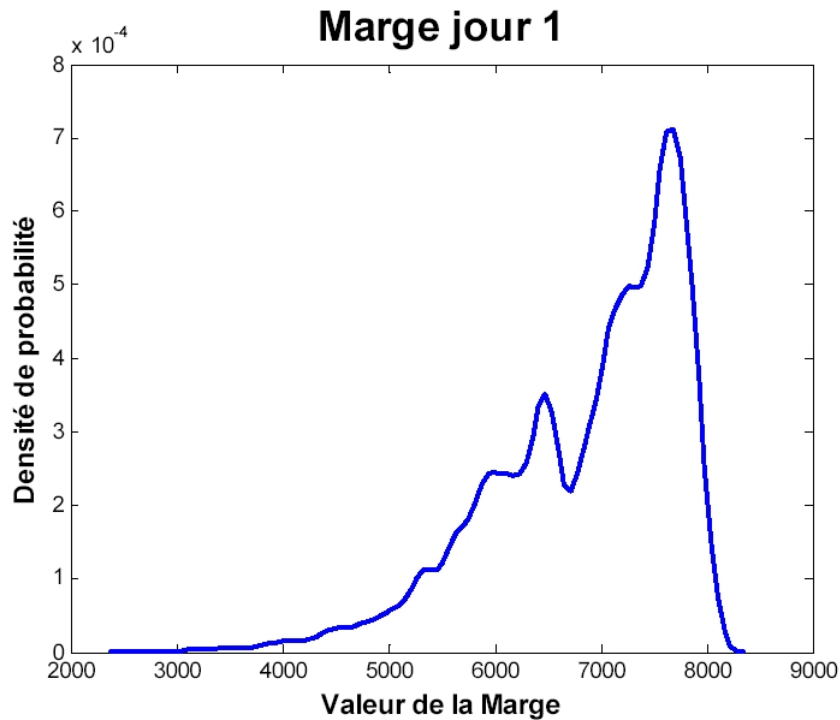
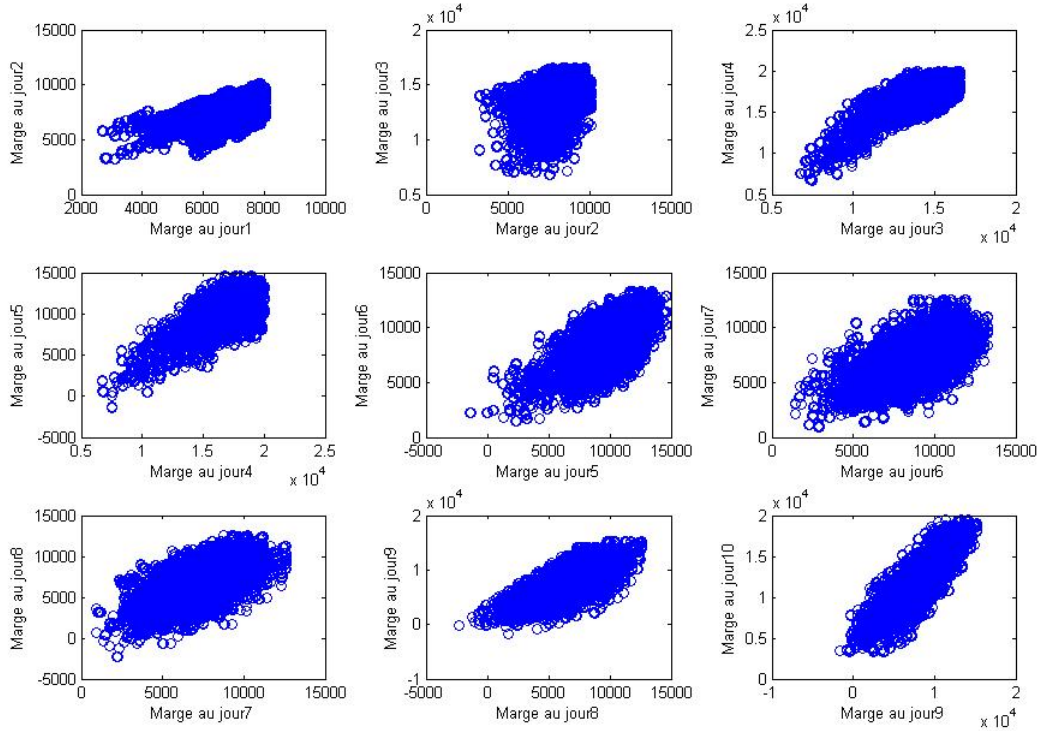


FIG. 7.4 – Un exemple de résultat de marge globale, exprimée en MW (Calcul avec 75000 scénarios pour le premier jour, semaine du 23 mars 2006)

7.2.7 Estimation de modèles ARMA

Il pourrait être intéressant d'avoir un modèle statistique pour représenter la marge de façon réaliste au sein de formulations de problèmes d'optimisation. Une étude statistique de la marge a été initiée en se concentrant sur des modèles ARMA, l'idée étant de pouvoir éventuellement intégrer de tels modèles dans des approches robustes. Cependant, ce processus stochastique est complexe. Sans procéder à une étude poussée, nous n'avons pu obtenir des résultats satisfaisants pour diverses raisons :

- Nous disposons de nombreux scénarios avec peu de pas de temps. Or, l'étude de séries temporelles est basée sur le contraire : une unique série avec un riche historique. La moyenne d'estimateurs ne convergeant pas vers un estimateur de qualité, cette difficulté est majeure. Il existe des méthodes particulières qui sont les méthodes de données de panels [12]. Il serait intéressant d'étudier ces approches. Une autre possibilité serait de générer des scénarios plus longs. Cependant, ces scénarios sont créés à partir du modèle opérationnel Apogée, adapté uniquement pour 144 pas de temps ;
- Problème de linéarité : si nous supposons un modèle autorégressif pour la marge, le tracé de M_t en fonction de M_{t+1} doit représenter une droite, justifiant la corrélation entre t et $t + 1$. Sur la figure ci-dessous, nous voyons que certains pas de temps sont plus représentés par des nuages de points très dispersés que par des droites (jours 2,5,6,7,...) :

FIG. 7.5 – M_t en fonction de M_{t+1}

- Pour remédier au problème des séries temporelles courtes, nous avons juxtaposé les scénarios pour ne créer qu'une unique série temporelle. Nous pouvons obtenir des résultats satisfaisants à condition que les modèles obtenus soient robustes à la taille de la série que nous créons. Or, nous avons remarqué qu'en fonction du nombre de scénarios juxtaposés, les modèles ARMA obtenus sont différents, et en particulier les ordres. Il est donc très difficile de caler un modèle satisfaisant.

7.3 Les prix de marché

7.3.1 Introduction

Le modèle à 2 facteurs et ses extensions constituent la classe de modèles la plus couramment utilisée pour la modélisation des prix à terme sur les marchés électriques, en particulier pour les applications en pricing. En effet ce modèle présente plusieurs avantages :

- Simplicité : il suffit de deux aléas élémentaires et trois paramètres pour décrire le marché électrique (marché spot et marché à terme). Le faible nombre d'aléas permet une simulation rapide (2 aléas déformant toute la courbe des prix à terme) ce qui constitue un réel intérêt pour les simulations de type Monte Carlo. Le faible nombre de paramètres qui ont une interprétation

simple permet à l'utilisateur d'appréhender facilement le modèle et de comprendre l'impact des paramètres sur les résultats.

- Modélisation jointe des prix Spot et des prix à terme : ce modèle permet d'explicitier le lien entre prix spot et prix à terme.
- Représentation sous forme d'une équation différentielle stochastique simple : cette représentation permet d'obtenir des formules de pricing explicites ou dans le cas d'options complexes des méthodes numériques relativement efficaces.
- Granularité horaire des prix à terme : cette caractéristique permet de valoriser des contrats offrant des options d'exercice au niveau de l'heure (comme les VPP²).

Il est clair que 3 paramètres ne suffisent pas pour représenter parfaitement le marché électrique dans son ensemble. Cependant, il faut noter que de façon générale plus les paramètres d'un modèle sont nombreux, plus le nombre de données nécessaires pour leur calage est important. Or les marchés électriques sont particulièrement non stationnaires ce qui rend inefficace l'utilisation d'un historique trop long. C'est pourquoi il peut parfois sembler plus sage d'utiliser un modèle plus fruste avec peu de paramètres mais qui sont réactualisés suffisamment fréquemment et judicieusement pour suivre la dynamique du marché. De plus, il faut noter que l'objectif de ce type de modèle n'est pas de représenter parfaitement le marché électrique dans son ensemble mais seulement de suivre les quelques caractéristiques du marché qui impactent l'application considérée (en l'occurrence, pour le type de modèles considéré ici, il s'agit le plus souvent du pricing).

7.3.2 Définitions

Définition 7.1 *Le terme **produit** désignera par la suite un type de contrat à terme caractérisé par sa durée de livraison et son préavis de livraison maximum (temps maximal avant la livraison). Nous raisonnons donc ici en “échéance relative” puisque la date de livraison n'est pas donnée. Sans plus de précision, nous faisons référence à des produits réellement cotés sur les marchés. Typiquement :*

- le 1-week ahead est coté au maximum 7 jours avant la livraison et correspond à une durée de livraison de 7 jours ;
- le 1-month ahead est coté au maximum 31 jours avant la livraison et correspond à une durée de livraison de 31 jours (environ) ;
- le 2-month ahead est coté au maximum 62 jours avant la livraison et correspond à une durée de livraison de 31 jours ;
- etc.

Pour un produit donné, nous disposons de plusieurs prix à terme. Ces prix à terme se distinguent soit par la date de cotation, soit par le préavis de livraison et la période de livraison (mois d'août ou juillet) ; ces deux dernières informations n'étant pas contenues dans la description du produit.

²Virtual Power Plant : centrales louées à d'autres producteurs par EDF pour favoriser l'ouverture à la concurrence

Définition 7.2 Le terme **produit unitaire** désignera par la suite un produit à terme de durée de livraison unitaire (en général l'unité est l'heure).

Définition 7.3 Un **contrat future** est, comme un contrat forward, un accord entre deux parties pour acheter ou vendre un actif donné à une date future pour un prix convenu. Contrairement aux contrats forward, les contrats futures sont négociés sur des marchés organisés. Les autorités de marché définissent des contrats standardisés pour assurer la liquidité.

Nous introduisons désormais quelques notions de mathématiques financières utilisées par la suite :

Définition 7.4 Un processus stochastique est une **martingale** par rapport à un ensemble d'information si son espérance en date t conditionnelle à l'information disponible en date $s < t$ est égale à la valeur du processus en date s , c'est-à-dire qu'un processus est une martingale si l'espérance conditionnelle de A_t en date s par rapport à la filtration \mathcal{F}_s est $A(s)$ (càd : $\mathbb{E}[A_t|\mathcal{F}_s] = A_s$).

Définition 7.5 Une des conséquences des hypothèses de non arbitrage (il n'existe aucune stratégie financière permettant, pour un coût initial nul, d'acquérir une richesse certaine dans une date future) et de complétude des marchés (tout flux à venir peut être répliqué exactement, et quel que soit l'état du monde, par un portefeuille d'autres actifs bien choisis) est l'existence et l'unicité à équivalence près d'une mesure de probabilité dite probabilité martingale ou "**probabilité risque-neutre**" telle que le processus de prix des actifs ayant une source de risque commune est une martingale sous cette probabilité. Cette probabilité peut s'interpréter comme celle qui régirait le processus de prix des sous-jacents de ces actifs si l'espérance du taux de rendement de ceux-ci était le taux d'intérêt sans risque (d'où le terme risque-neutre : aucune prime n'est attribuée à la prise de risque).

Définition 7.6 Nous appelons **mouvement brownien** un processus stochastique $(X_t)_{t \geq 0}$ à valeurs réelles, à accroissements indépendants et stationnaires et dont les trajectoires sont continues. Ce qui signifie que :

- continuité : la fonction $s \mapsto X_s(\omega)$ est une fonction continue presque sûrement ;
- indépendance des accroissements : si $s \leq t$, $X_s - X_t$ est indépendant de la tribu $\mathcal{F}_s = \sigma(X_u, u \leq s)$;
- stationnarité des accroissements : la loi de $X_t - X_s$ est identique à celle de $X_{t-s} - X_0$.

7.3.3 Du modèle factoriel unitaire au modèle factoriel par produit

Le modèle à 2 facteurs vise à modéliser la dynamique des prix à terme correspondant à des périodes de livraison unitaires (horaires) et avec un temps de cotation supposé continu. Dans la réalité, les produits cotés n'ont pas tous la même durée de livraison (heure, semaine, mois, etc.) et les instants de cotations sont discrets (tous les jours). Le modèle proposé ne correspond donc pas directement aux données de marché observées. Pour simplifier la phase de calage du "modèle factoriel unitaire en temps continu" à partir des données est nécessaire. Pour ce qui concerne le caractère échantillonné des dates de cotation, la solution est de discrétiser le modèle en temps.

Pour ce qui concerne le caractère non unitaire des périodes de livraison, nous déduisons du modèle factoriel unitaire un modèle factoriel par produit (non unitaire). Nous obtenons ainsi un modèle de prix qui correspond aux données de marché observées. Cette approche nécessite tout de même une approximation pour être utilisable pour le calage.

Le modèle à 2 facteurs unitaire en temps continu

Nous rappelons ici la description du modèle à 2 facteurs unitaire en temps continu.

Nous notons $F(t, T)$ le prix future à l'instant t pour une livraison de 1 MWh à la date T sur une durée de livraison unitaire. L'unité de temps choisie est usuellement l'heure si nous voulons rester cohérents avec la définition du prix spot $S_t = F(t, t)$ qui correspond effectivement à une livraison horaire. Dans la suite, nous adopterons par défaut ce choix d'unité sauf indication contraire. Nous supposons que sous la probabilité risque neutre, les prix futures vérifient l'équation différentielle (EDS) du modèle à 2 facteurs [13], c'est-à-dire :

$$dF(t, T) = F(t, T) \left(\beta_s(t) e^{-\alpha(T-t)} d\tilde{W}_t^s + \beta_l(t) d\tilde{W}_t^l \right), \quad (7.9)$$

où la volatilité court terme β_s et la volatilité long terme β_l sont des fonctions déterministes du temps, α est le coefficient de retour à la moyenne supposé constant. $(d\tilde{W}_t^s, d\tilde{W}_t^l)_{t \geq 0}$ est un mouvement brownien (en dimension 2) standard sous la probabilité risque neutre. Sous la probabilité réelle l'EDS vérifiée par les futures peut se mettre sous la forme générale :

$$dF(t, T) = F(t, T) \left(\left(\gamma_l(t) + \gamma_s(t) e^{-\alpha(T-t)} \right) dt + \beta_s(t) e^{-\alpha(T-t)} dW_t^s + \beta_l(t) dW_t^l \right), \quad (7.10)$$

où $\gamma_l(t) + \gamma_s(t) e^{-\alpha(T-t)}$ est la prime de risque (γ_s et γ_l sont des fonctions du temps) et $(W_t^s, W_t^l)_{t \geq 0}$ est un mouvement brownien (en dimension 2) standard sous la probabilité réelle. Par intégration (cf. [13]), nous obtenons :

$$F(t, T) = F(0, T) \exp \left(\left(m(t, T) - \frac{1}{2} \sigma^2(t, T) \right) t + \int_0^t \beta_s(u) e^{-\alpha(T-u)} dW_u^s + \beta_l(u) dW_u^l \right), \quad (7.11)$$

avec

$$m(t, T) = \frac{1}{t} \int_0^t \gamma_l(u) + \gamma_s(u) e^{-\alpha(T-u)} du \text{ et } \sigma^2(t, T) = \frac{1}{t} \int_0^t (\beta_l^2(u) + \beta_s^2(u) e^{-2\alpha(T-u)}) du.$$

$m(t, T)$ et $\sigma(t, T)$ correspondent respectivement à la dérive (tendance) et à la volatilité moyenne du prix à terme entre les instants 0 et t pour livraison en T .

Remarque 7.2 Nous avons supposé nulle la corrélation entre facteur court terme et facteur long terme. La prise en compte de cette corrélation rajouterait des termes croisés et un paramètre supplémentaire à caler, mais ne compliquerait pas fondamentalement les résultats.

Le modèle à 2 facteurs “par produit” en temps continu

Dans cette sous-section, nous écrivons le prix future correspondant à un produit non unitaire à partir de prix futures unitaires. Nous en déduisons un modèle de dynamique des prix futures non unitaire (nous parlons de modèle par produit) en temps continu à partir du modèle unitaire en temps continu.

Comme nous disposons d'historiques de prix correspondant à des périodes de livraison non unitaires allant de l'heure (pour le prix spot) à l'année (pour les year-ahead), nous nous intéressons au modèle vérifié par les produits directement.

Nous notons $F(t, T, \theta)$ le prix future coté en t pour une livraison entre T et $T + \theta$. Par absence d'opportunité d'arbitrage, nous avons :

$$F(t, T, \theta) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} F(t, T + i). \quad (7.12)$$

Par différentiation, nous obtenons :

$$\begin{aligned} dF(t, T, \theta) &= \frac{1}{\theta} \sum_{i=0}^{\theta-1} dF(t, T + i), \\ &= \frac{1}{\theta} \sum_{i=0}^{\theta-1} \left[\left(\gamma_l(t) + \gamma_s(t) e^{-\alpha(T+i-t)} \right) dt + \beta_s(t) e^{-\alpha(T+i-t)} dW_t^s + \beta_l(t) dW_t^l \right] F(t, T + i). \end{aligned}$$

L'objectif est de trouver un modèle de dynamique aléatoire vérifié par les prix par produits $F(t, T, \theta)$ qui ne fasse plus intervenir les prix unitaires $F(t, T + i)$ comme dans le terme de droite de l'équation ci-dessus. Pour cela, il suffit d'exprimer les prix “unitaires” $F(t, T + i)$ en fonction des prix “par produit” $F(t, T, \theta)$. Une première approximation très grossière consiste à dire

$$F(t, T + i) \approx F(t, T, \theta), \quad \forall i = 0, \dots, \theta - 1,$$

ce qui suppose que les prix futures “unitaires” sont égaux et ont la même contribution dans le produit. Cette approximation est contradictoire avec une saisonnalité des prix observée et semble a priori trop grossière. Une meilleure approximation demande une pondération des futures “unitaires”. Cette dernière est faite par la prise en compte des facteurs de forme.

Introduction des facteurs de forme

Pour prendre en compte la fluctuation des prix à l'intérieur d'une période de livraison (un prix future pour livraison le lundi à 8H00 est probablement plus élevé qu'un prix future pour livraison le dimanche à 0H00), nous introduisons les coefficients

$$\lambda_i^{t,T,\theta} = \frac{F(t, T+i)}{F(t, T, \theta)}, \quad \forall i = 0, \dots, \theta - 1. \quad (7.13)$$

Nous notons que les coefficients $\lambda_i^{t,T,\theta}$ ainsi définis peuvent s'interpréter comme des facteurs de forme (pondérant le prix de l'heure i par rapport au prix moyen sur la période de livraison $[T, T + \theta]$). Avec cette définition, la somme des coefficients correspondant à une période de livraison donnée est égale à la durée de livraison : $\sum_{i=0}^{\theta-1} \lambda_i^{t,T,\theta} = \theta$. Nous notons de plus que ces facteurs de forme dépendent de la date de cotation des produits considérés t et qu'ils sont aléatoires. Avec cette notation, nous obtenons :

$$dF(t, T, \theta) = \frac{1}{\theta} F(t, T, \theta) \sum_{i=0}^{\theta-1} \lambda_i^{t,T,\theta} \left[(\gamma_l(t) + \gamma_s(t) e^{-\alpha(T+i-t)}) dt + \beta_s(t) e^{-\alpha(T+i-t)} dW_t^s + \beta_l(t) dW_t^l \right].$$

En divisant par $F(t, T, \theta)$, nous avons

$$\frac{dF(t, T, \theta)}{F(t, T, \theta)} = [\gamma_l(t) + \gamma_s(t) \frac{1}{\theta} \sum_{i=1}^{\theta-1} \lambda_i^{t,T,\theta} e^{-\alpha(T+i-t)}] dt + \beta_s(t) \frac{1}{\theta} \sum_{i=1}^{\theta-1} \lambda_i^{t,T,\theta} e^{-\alpha(T+i-t)} dW_t^s + \beta_l(t) dW_t^l. \quad (7.14)$$

Nous posons

$$\psi(t, T, \theta) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} \lambda_i^{t,T,\theta} e^{-\alpha i}, \quad (7.15)$$

la somme pondérée des facteurs de forme sur la période de livraison. Nous notons que dans $\psi(t, T, \theta)$ la contribution des facteurs de forme de fin de période de livraison est négligeable (exponentiellement) par rapport aux facteurs de forme de début de livraison. Avec cette notation, nous obtenons finalement l'EDS suivante :

$$\frac{dF(t, T, \theta)}{F(t, T, \theta)} = \gamma_l(t) dt + \gamma_s(t) e^{-\alpha(T-t)} \psi(t, T, \theta) dt + \beta_s(t) e^{-\alpha(T-t)} \psi(t, T, \theta) dW_t^s + \beta_l(t) dW_t^l. \quad (7.16)$$

Nous notons que le terme $\psi(t, T, \theta)$ dépend du temps. Il est aléatoire puisqu'il dépend des prix à terme unitaires à l'instant t . Cela implique que la dynamique des prix à terme non unitaires n'est ni markovienne ni gaussienne.

7.3.4 Discrétisation en temps du modèle à 2 facteurs “par produit”

Dans cette section, l’objectif est de discrétiser le modèle (7.16) pour obtenir finalement un modèle “par produit” en temps discret qui correspond à la forme des données observées. Nous pourrions faire une discrétisation grossière basée sur l’approximation d’Euler conduisant à des rendements proportionnels. Mais, même si cette approche est très utilisée pour sa simplicité, les approximations peuvent être importantes et pour cette raison nous préférons utiliser une discrétisation par rendements logarithmiques proposée ci-dessous.

Nous notons $F_n(t, T)$ le prix future d’1 MWh coté à l’instant t_n pour une livraison entre T et $T + \theta$, nous avons donc $F_n(T, \theta) = F(t_n, T, \theta)$. Nous avons la relation suivante entre prix par produit et prix “horaire” :

$$F_n(T, \theta) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} F_n(T + i, 1) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} F(t_n, T + i).$$

Nous discrétisons l’EDS (7.16) en posant que le pas de discrétisation $t_{n+1} - t_n = \delta_t$ (comme les cotations des prix futures sont journalières, le pas de discrétisation δ_t vaut en général un jour).

Si nous raisonnons sur les rendements logarithmiques, en utilisant le lemme d’Ito à la fonction “log” entre les instants t_n et t_{n+1} , avec la seule approximation que les coefficients $\gamma_l(t)$, $\gamma_s(t)$, $\beta_s(t)$, $\beta_l(t)$ et $\psi(t, T, \theta)$ restent constants entre t_n et t_{n+1} , nous obtenons le modèle suivant :

$$R_n(T, \theta) = \log \left(\frac{F_{n+1}(T, \theta)}{F_n(T, \theta)} \right) \quad (7.17)$$

$$= \gamma_l(t_n) \delta_t + \gamma_s(t_n) \psi(t_n, T, \theta) e^{-\alpha(T-t_n)} \mu(\delta_t) - \frac{1}{2} \sigma_n^2(T, \theta) \quad (7.18)$$

$$+ \beta_s(t_n) e^{-\alpha(T-t_n)} \psi(t_n, T, \theta) \sqrt{\nu(\delta_t)} \epsilon_n^s + \beta_l(t_n) \sqrt{\delta_t} \epsilon_n^l \quad (7.19)$$

où

$$\begin{cases} \mu(\delta t) = \frac{e^{\alpha \delta t} - 1}{\alpha} \\ \nu(\delta t) = \frac{e^{2\alpha \delta t} - 1}{2\alpha} \\ \sigma_n^2(T, \theta) = \beta_s^2(t_n) \psi^2(t_n, T, \theta) e^{-2\alpha(T-t_n)} \nu(\delta_t) + \beta_l^2(t_n) \delta_t. \end{cases}$$

Nous précisons que les aléas intervenant dans le modèle (7.17) sont donnés par :

$$\begin{aligned} \epsilon_n^s &= \frac{1}{\sqrt{\nu(\delta t)}} \int_{t_n}^{t_{n+1}} e^{-\alpha(t_n-u)} dW_u^s \\ \epsilon_n^l &= \frac{1}{\sqrt{\nu(\delta t)}} (W_{t_{n+1}}^l - W_{t_n}^l). \end{aligned}$$

Ces deux modèles supposent une structure de volatilité par terme exponentiellement décroissante lorsque nous nous rapprochons de la maturité du produit et une volatilité qui grossièrement est inversement proportionnelle à la durée de livraison du fait du coefficient $\psi(t_n, T, \theta)$.

7.3.5 Approximation sur les facteurs de forme

Dans cette section, nous nous intéressons particulièrement à la somme pondérée des facteurs de forme $\psi(t_n, T, \theta)$. Pour pouvoir appliquer des méthodes d'estimation statistiques simples et efficaces, nous préférons que le modèle discrétisé soit gaussien et markovien. Or le modèle (7.17) n'est ni gaussien ni markovien. En effet, la dynamique fait intervenir le terme $\psi(t_n, T, \theta)$ et donc $\lambda_i^{t_n, T, \theta}$. Comme ce terme dépend des prix unitaires $F(t_n, T + i)$ et non uniquement du prix par produit $F_n(T, \theta)$, nous remarquons que les prix $F_n(T, \theta)$ ne sont pas gaussiens. De plus, comme $\psi(t_n, T, \theta)$ est aléatoire et intervient de façon multiplicative sur l'aléa gaussien, nous concluons que les rendements des prix $F_n(T, \theta)$ ne sont pas gaussiens. Pour approcher le modèle discrétisé par un modèle markovien et gaussien, nous sommes amenés à faire l'approximation que les facteurs de forme sont constants par rapport à la date de cotation t , $\lambda_i^{t, T, \theta} = \lambda_i^{T, \theta}$. Autrement dit, nous supposons que pour tout j entre 0 et θ , nous avons :

$$\lambda_j^{t, T, \theta} = \frac{F(t, T + j)}{\frac{1}{\theta} \sum_{i=0}^{\theta-1} F(t, T + i)} = \frac{F(t_0, T + j)}{\frac{1}{\theta} \sum_{i=0}^{\theta-1} F(t_0, T + i)} = \lambda_j^{T, \theta}, \quad \forall t.$$

Avec cette hypothèse, nous obtenons que $\psi(t, T, \theta) = \psi(T, \theta)$ est déterministe et indépendant de l'instant de cotation,

$$\psi(t, T, \theta) = \psi(T, \theta) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} \lambda_i^{T, \theta} e^{-\alpha i}.$$

En pratique, nous supposons tous les facteurs de forme égaux à 1 et donc

$$\psi(t, T, \theta) = \psi(\theta) = \frac{1}{\theta} \frac{1 - e^{-\alpha \theta}}{1 - e^{-\alpha}}.$$

Sous ces nouvelles hypothèses, le modèle devient :

$$\begin{aligned} R_n(T, \theta) &= \gamma_l(t_n) \delta t + \gamma_s(t_n) \psi(T, \theta) e^{-\alpha(T-t_n)} \mu(\delta t) - \frac{1}{2} \sigma_n^2(T, \theta) + \beta_s(t_n) e^{-\alpha(T-t_n)} \psi(T, \theta) \sqrt{\nu(\delta t)} \epsilon_n^s \\ &+ \beta_l(t_n) \sqrt{\delta t} \epsilon_n^l, \end{aligned}$$

où :

$$\begin{cases} \mu(\delta t) = \frac{e^{\alpha \delta t} - 1}{\alpha} \\ \nu(\delta t) = \frac{e^{2\alpha \delta t} - 1}{2\alpha} \\ \sigma_n^2(T, \theta) = \beta_s^2(t_n) \psi^2(T, \theta) e^{-2\alpha(T-t_n)} \nu(\delta t) + \beta_l^2(t_n) \delta t. \end{cases}$$

7.3.6 Création de scénarios

En pratique, la génération de scénarios est assez simple. Une fois les rendements calculés, nous utilisons l'équation (7.17) pour reconstruire les prix puisque :

$$F_{n+1}(T, \theta) = F_n(T, \theta) \exp(R_n(T, \theta))$$

Les paramètres utilisés sont issus de Energy Data Base³. Pour les exemples numériques considérés dans cette thèse, nous avons utilisé les paramètres issus de la période du 01 janvier 2006 au 01 octobre 2006 suivants :

$$\begin{aligned}\beta^l &= \frac{0.2749}{100 * \sqrt{365}} \\ \beta^s &= \frac{65.04}{100 * \sqrt{365}} \\ \alpha &= \frac{7.91}{365}\end{aligned}$$

Bibliographie

- [1] B. Beuzamy. Méthodes Probabilistes pour l'Etude des Phénomènes Réels. Technical report, Société de Calcul Mathématique S.A., 2004.
- [2] B. Beuzamy. Gestion Hebdomadaire de la Production : Démarche Méthodologique de la Prise en Compte d'Aléas dans le Modèle Apogée. Technical report, Société de Calcul Mathématique S.A., Etude EDF-SCM n°8610-AAR-4300023230, 2005.
- [3] H. Blondel and A. Renaud. Décomposition par les Prix et Lagrangien Augmenté : une Association Efficace pour le Traitement des Problèmes de Production. Bulletin de la Direction des Etudes et Recherches d'Electricité de France, Série C, N°4.
- [4] A.W. Bowman and A. Azzalini. Applied Smoothing Techniques for Data Analysis : the Kernel Approach with S-Plus Illustrations. *Oxford University Press*, 1997.
- [5] G. Cohen and D.L. Zhu. Decomposition-Coordination Methods in Large-Scale Optimization Problems. The Nondifferentiable Case and the Use of Augmented Lagrangians. *Advance in Large Scale Systems, Theory and Applications*, Vol. 1, CRUZ J.B. Editor, JAI Press, Inc., USA, 1983.
- [6] P. Deheuvels and P. Hominal. Estimation Automatique de la Densité. *Revue de Statistique Appliquée*, tome 28, n°1 :25–55, 1980.
- [7] A. Merlin and P. Sandrin. A new Method for Unit Commitment at Electricité de France. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, N°5 :1218–1225, 1983.
- [8] Moché. Estimation de la Densité par les Histogrammes et les Estimateurs à Noyau. Technical report, CRDP, Lille, 2004.

³Base de Données stockant les historiques de prix de chaque contrat

- [9] N. Oudjane. Etude des Méthodes de Calage du Modèle à 2 Facteurs pour le Pricing et la Gestion des Risques. Technical report, EDF R&D, Dept OSIRIS, HR-33/04/009/A, 1 av. Général de Gaulle, F-92141 Clamart, 2004.
- [10] A. Prékopa. *Stochastic Programming*. Kluwer, Dordrecht, Boston, 1995.
- [11] A. Renaud. *Algorithmes de Régularisation et Décomposition pour les Problèmes Variationnels Monotones*. PhD thesis, Ecole Nationale des Mines de Paris, Spécialité Mathématiques et Automatique, 1993.
- [12] P. Sevestre. *Econométrie des Données de Panel*. Dunod, 2002.
- [13] F. Turbault and X. Warin. Modélisation des Futures et Calcul Numérique d’Option. Technical report, EDF R&D, Dept MMN, HI-72/2000/009/A, 2000.
- [14] G. Vignal. Modèle APOGEE. Vers une Nouvelle Formulation des Sous-Problèmes Hydrauliques. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/003/A, 2004.
- [15] W. Wertz and B. Schneider. Statistical Density Estimation : a Bibliography. *International Statistical Review*, 47 :155–175, 1979.
- [16] R. Zorgati. Étude Prospective sur une Méthodologie de Prise en Compte d’Aléas et d’Analyse des Risques Associés dans le Modèle Déterministe APOGEE 12 Jours - Cahier des Charges. Technical report, EDF R&D, Dept OSIRIS, 2004.
- [17] R. Zorgati, R. Apparigliato, and J.-M. Luel. Couverture et Arbitrage sur les Marchés en Gestion Hebdomadaire de la Production Electrique. Partie I : Marge de Production et Risque Physique à 1%. Technical report, EDF R&D, Dept OSIRIS, H-R37-2007-01772-FR, 1 av. Général de Gaulle, F-92141 Clamart, 2007.
- [18] R. Zorgati and P. Thomas. Formation au Modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d’une Vallée sur Critère Economique. I. Formulation. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/015/A, 2004.
- [19] R. Zorgati and P. Thomas. Formation au Modèle APOGEE-12 Jours. Initiation au Placement Optimal de la Production Hydraulique d’une Vallée sur Critère Economique. II. Résolution. Technical report, EDF R&D, Dept OSIRIS, HR-37/04/016/A, 2004.
- [20] R. Zorgati, W. Van Ackooij, and R. Apparigliato. Supply Shortage Hedging : Estimating the Electrical Power Margin for Optimizing Financial and Physical Assets with Chance-Constrained Programming. *submitted to IEEE Transaction on Power Systems*, March 26th, 2008.
- [21] R. Zorgati, W. Van Ackooij, J.-M. Luel, P. Thomas, M. Uchanski, and K. Shea. LEA : A Cluster Intensive Simulation Software For Unit Commitment. *Proceedings of the 7th LCI High Performance Computing Conference, Norman, Oklahoma, (USA)*, May 02-04 2006.

Quatrième partie

Conclusions et perspectives

Chapitre 8

Conclusions

Le thème de cette thèse est la gestion des risques physiques à l'horizon hebdomadaire. Cette thématique étant relativement vaste, deux problématiques industrielles sont particulièrement intéressantes : la gestion court-terme de la production hydro-électrique et l'optimisation de la couverture contre le risque de défaillance physique. Elles permettent de contribuer à l'amélioration des deux rôles de l'horizon hebdomadaire : simuler l'état du parc sur les deux prochaines semaines et prendre des décisions de sécurité pour couvrir le risque de défaillance physique. Ces problèmes d'optimisation sont d'un enjeu économique important, et fortement incertains car sujets à des aléas très variables (hydraulicité, indisponibilité, prix de marché,...). Les techniques d'optimisation appliquées dans nos travaux satisfont des objectifs que nous nous sommes fixés : elles répondent aux problèmes par des solutions pertinentes, facilement interprétables pour des opérationnels, et elles conservent des temps de calcul raisonnables. Pour cela, il a donc été nécessaire de maîtriser les cadres théoriques de chacune, les évaluer et parvenir à les appliquer à nos problématiques. Nous nous sommes ainsi intéressés à des approches présentes dans la littérature mais jusqu'à présent pas appliquées à notre connaissance dans le secteur énergétique.

La première approche utilisée est l'optimisation robuste. Elle a été appliquée dans cette thèse à la gestion stochastique des vallées hydrauliques. Cette étude a montré que nous pouvons réduire drastiquement (voire supprimer) les violations de contraintes sur les bornes des stocks hydrauliques sans dégradation importante des coûts. De plus, ce résultat favorable peut s'obtenir sans accroître sérieusement la complexité du modèle et de la méthode de résolution. Enfin, les modèles intégrant des règles de décision semblent particulièrement prometteurs (réduction de violations quasi-similaires) bien que leur taille soit plus importante.

L'autre approche est la programmation stochastique avec règles de décision constantes par morceaux (SPSDR). Elle se base sur la méthodologie de la programmation stochastique classique sur arbre d'aléas mais cherche à contourner le problème de l'explosion de l'arbre en supprimant la contrainte de raffinement. La décision finale peut également être améliorée en utilisant l'avis de plusieurs experts. Nous avons par ailleurs proposé des extensions possibles à cette heuristique : aléa multidimensionnel et ajout de contraintes sur les états. Les tests ont été mis en place sur une version simplifiée du problème d'optimisation de la couverture physique (petite taille, variables continues). Les résultats sont prometteurs puisqu'ils satisfont les critères désirés : satisfaction

du critère de risque en probabilité et minimisation des coûts. Les bénéfices économiques retirés sont même supérieurs à des approches statiques. Cependant, les nombreux paramètres qu'il convient de régler et l'absence de formule permettant de prédire l'impact de ces paramètres sur la violation des contraintes imposent d'avoir recours à une approche basée sur de nombreuses expérimentations.

Des résultats théoriques ne sont pas toujours disponibles pour s'assurer de la convergence vers une solution (en particulier pour SPSDR). C'est pourquoi elles ont été validées par de la simulation numérique. L'usage de la simulation est un aspect important dans cette thèse et une pratique relativement rare dans le domaine de l'optimisation stochastique. Nous nous sommes notamment aperçus que, bien que très riche, la littérature de l'optimisation dans l'incertain ne validait que rarement les stratégies obtenues en optimisation. Or, ces stratégies sont obtenues à l'aide d'hypothèses particulières sur les aléas (discrétisation par des scénarios ou des arbres,...) et, finalement, le gestionnaire est nullement assuré que la stratégie obtenue soit aussi satisfaisante face à des scénarios inédits que lors de la phase d'optimisation. Cette phase de simulation est relativement expérimentale, mais elle donne une indication sur la qualité de la solution, information nécessaire dans un cadre industriel.

D'autre part, une contribution importante est la modélisation du problème de gestion de la couverture physique. Au delà des approches de résolution, il est très important d'avoir analysé ce problème à l'horizon hebdomadaire de façon globale (couplant la gestion de production, l'optimisation de la couverture et la politique de risque) et d'avoir proposé une formulation cohérente, capable de fournir une politique de couverture efficace. Elle permet aussi de valider la procédure actuelle des gestionnaires tout en la justifiant mathématiquement et en proposant des améliorations. D'ailleurs, une meilleure prise en compte de la topologie des vallées hydrauliques (permettant d'avoir une capacité maximale de production hydraulique réalisable) et l'intégration de modélisations plus réalistes des aléas dans le calcul des scénarios de marge de production physique sont des plus values importantes.

Comme nous venons brièvement de le rappeler, cette thèse CIFRE a permis d'appliquer, d'analyser et adapter des approches particulières à des problématiques industrielles, tout en apportant une plus value métier à l'aide de nouvelles modélisations. Cette thèse été l'occasion d'explorer de nouvelles approches et de nouvelles méthodes pour aborder des problématiques du secteur énergétique. Nous pensons qu'elle ouvre ainsi des perspectives intéressantes et pose des questions stimulantes. Certaines d'entre elles ont été abordées dans le dernier chapitre de cette thèse.

Chapitre 9

Quelques perspectives pour l'après thèse

Sommaire

9.1	Perspectives sur la gestion robuste d'une vallée hydraulique	240
9.1.1	Pistes sur l'application de l'optimisation robuste	240
9.1.2	Pistes sur la modélisation du problème	241
9.2	Perspectives sur la gestion des risques de défaillance physique . . .	241
9.2.1	Le calcul de la marge	242
9.2.2	Boucle ouverte : calcul de marge intégré au problème de couverture . .	243
9.2.3	Boucle fermée	250

Dans ce dernier chapitre, nous proposons quelques extensions aux travaux effectués tout au long de cette thèse. Ce sont des pistes qui nous semblent particulièrement intéressantes pour la suite. Elles sont basées, soit sur de simples intuitions, soit à partir de modèles donnant des premiers éléments de réponse prometteurs, mais pas suffisamment aboutis faute de temps.

9.1 Perspectives sur la gestion robuste d'une vallée hydraulique

Sur cette étude, nous pouvons distinguer deux principaux axes d'amélioration : une meilleure application de la méthodologie robuste au problème et une complexification de la modélisation de la vallée.

9.1.1 Pistes sur l'application de l'optimisation robuste

Dans notre synthèse sur l'optimisation robuste (voir chapitre 3), nous avons décrit des extensions capables d'améliorer le modèle robuste. La plus intéressante de toutes dans le cadre de la gestion hydraulique semble être la **robustesse étendue**. Nous rappelons qu'il s'agit de contrôler les violations de contrainte lorsque l'aléa se réalise en dehors de l'ensemble d'incertitude. La violation est bornée par une fonction proportionnelle à la distance de l'aléa à l'ensemble d'incertitude. Dans des périodes où la variabilité des apports est particulièrement importante, cette technique permettrait au modèle de contrôler les violations sans avoir à considérer des ensembles d'incertitude plus grands que pour une "période normale". Le modèle robuste peut également être amélioré par l'introduction de **règles de décision avec paramètres de recours fixes**. Elles permettent d'appliquer des stratégies différentes selon que l'aléa est plus ou moins élevé par rapport à la moyenne.

Nous pouvons également nous intéresser à la construction des ensembles de décision. Au lieu d'approcher une boule ℓ_2 par un polyèdre, nous pouvons en fait **conserver directement cette boule ℓ_2 en tant qu'ensemble d'incertitude**. En effet, l'approximation simpliste consistant à construire un polyèdre en intersectant les normes usuelles ℓ_1 et ℓ_∞ induit une erreur importante lorsque la dimension du vecteur d'aléas est grande. Comme le montre le lemme 3.1 du chapitre 3, l'équivalent robuste devient alors un problème d'optimisation conique-quadratique, problème de complexité proche de la programmation linéaire. Nous effectuons tout de même un saut apparent dans la difficulté puisqu'il est nécessaire d'avoir de nouveaux outils adéquats (solveurs).

Enfin, nous pouvons aussi étudier l'**utilisation de nouvelles règles de décision**, plus complexes que celles utilisées dans cette étude. Les règles utilisées ont toujours été linéaires, dépendantes des réalisations passées de l'aléa. Quitte à accroître la difficulté du problème, nous pourrions quitter le cadre affine des règles en l'aléa même si c'est à l'encontre de notre objectif de conserver des équivalents robustes de complexité proche du problème initial.

9.1.2 Pistes sur la modélisation du problème

La première piste d'évolution est la transformation du jeu d'essais. Il serait tout à fait raisonnable de tester ces approches sur des vallées différentes pour confirmer nos résultats. Nous pouvons considérer par exemple **des vallées plus grandes ou introduire des capacités de pompage**.

La deuxième piste est l'**intégration de variables de décision entières**. Dans le cadre des modèles sans ajustement linéaire, il semble évident que le passage des variables continues à un caractère discret s'effectue sans problème (hormis les temps de calculs qui devraient devenir plus élevés). Dans le cadre de modèles avec ajustements linéaires, le problème est plus complexe. Le premier jour peut être traité à part puisque les commandes ne peuvent être des règles (nous ne pouvons pas faire dépendre les commandes d'un aléa passé qui n'existe pas encore). Pour les autres jours, cela semble impossible en optimisation. Même si les coefficients de la règle sont imposés entiers, il est nécessaire que les aléas se réalisent en tant que variable aléatoire entière, chose impossible en pratique. Donc rien nous assurerait que la commande finale soit entière. Par contre, nous pourrions imaginer une phase de simulation qui transformerait les variables continues en des variables discrètes à l'aide par exemple d'heuristiques adaptées.

Le dernier point, et non le moindre, est l'**intégration d'un aléa supplémentaire** : l'aléa de prix. Jusqu'à présent, l'aléa était uniquement en second membre des contraintes. Or, les prix de marché sont en produit avec les variables. Cela peut poser un problème, notamment si nous intégrons des ajustements linéaires, puisque nous nous retrouvons de fait avec des produits d'aléas. La résolution d'un tel problème reste ouverte mais il semblerait que la programmation conique quadratique pourrait résoudre ce problème efficacement. Nous pourrions ainsi imaginer une gestion hydraulique où les stocks seraient mis face aux marchés : si le prix est intéressant, nous produisons de l'électricité pour le revendre, sinon nous conservons l'eau. D'autres possibilités pour résoudre le problème avec un prix aléatoire sont imaginables. Nous pouvons ainsi nous intéresser au couplage d'approches comme l'optimisation robuste (pour l'aléa d'hydraulicité) et de la programmation stochastique (pour les prix). En chaque nœud d'un arbre, c'est à dire pour chaque réalisation d'un aléa de prix, nous nous assurerions que la solution obtenue serait robuste à l'aléa d'hydraulicité. Ces modèles seraient beaucoup plus complexes, de très grande taille (selon la taille de l'arbre), mais pourraient s'avérer très intéressants.

9.2 Perspectives sur la gestion des risques de défaillance physique

Dans cette deuxième partie de thèse, nous avons initié une nouvelle problématique et commencé à proposer des réponses. De nombreuses pistes se sont ouvertes et nous semblent particulièrement intéressantes. Dans cette section, nous allons donc en présenter quelques unes.

9.2.1 Le calcul de la marge

Même si l'approche proposée dans le chapitre 7 améliore sans aucun doute le calcul de la marge de production, il reste encore des points possibles d'amélioration :

- Pour le calcul de la capacité maximale de production thermique, toutes les centrales disponibles sont supposées pouvoir produire à leur maximum à l'heure de pointe. Or, une centrale en fonctionnement doit satisfaire des contraintes techniques particulières comme des paliers de production, des changements de paliers progressifs, des courbes de démarrage,... Une centrale n'est donc pas forcément disponible à son maximum à l'heure de pointe. Pour modéliser ces contraintes, il est nécessaire de résoudre, à l'image de la capacité maximale hydraulique, **le problème d'optimisation du planning de production thermique sous aléa d'indisponibilité**. A l'horizon hebdomadaire, ce problème est résolu par programmation dynamique à l'aide d'un graphe d'états prédéfini et fonction de l'état initial du parc de centrales. Nous pouvons imaginer pouvoir résoudre ce problème par programmation dynamique stochastique. Par contre, les temps de calculs risquent de devenir assez rapidement très élevés ;
- Pour le calcul de la capacité maximale de production hydraulique, la situation est plus complexe. Il y a deux limites clairement détectées : la chronique de prix fixe, construite en augmentant arbitrairement les prix aux heures de pointe et la non prise en compte de l'histoire de la marge hydraulique. Ce problème a déjà été discuté et est basé sur le fait que la capacité hydraulique en un jour dépend des quantités d'eau qui ont été utilisées ou stockées. Comme cette information est contenue dans le calcul du volume, l'histoire ne dépend que du dernier état du système : c'est donc une chaîne de Markov. Un moyen de prendre en compte cette histoire est de discrétiser l'état du problème (le volume de chaque réservoir) et de chercher la trajectoire optimale sur les J jours. Nous pouvons soit conserver un prix fixe soit considérer que le prix est une variable aléatoire. L'approche classique pour résoudre ce type de problèmes est la programmation dynamique stochastique. Cependant, la programmation dynamique (stochastique ou non) est sujette à la "malédiction de la dimension", c'est à dire qu'il est quasi impossible de résoudre un problème avec un nombre de stocks élevé (en pratique, autour de cinq selon le niveau de discrétisation). Or, le problème hydraulique contient une quarantaine de vallées hydrauliques. Même si nous agrégeons les lacs d'une vallée voire même des vallées entre elles, le nombre de stocks demeurerait trop important. Sauf si, à la manière du moyen terme, toutes les quantités d'eau sont agrégées en un "lac France". Mais nous perdrons la finesse de modélisation nécessaire à la modélisation hebdomadaire. La précision gagnée sur la prise en compte de l'histoire serait compensée par cette perte de modélisation. Par contre, une approche possible est la **programmation dynamique stochastique duale** (voir chapitre 2). Cette méthode effectue des approximations des fonctions de Bellman via des méthodes de coupe. A l'inverse de la programmation dynamique classique, elle peut être utilisée pour un grand nombre de stocks. Cependant, elle propage une erreur d'approximation d'un pas de temps sur l'autre qui peut devenir importante si le nombre de périodes est trop importante. Or, à l'horizon hebdomadaire, le nombre de périodes est faible (quinze tout au plus). Cette approche semble donc parfaitement intéressante dans notre cas. Par contre, il est bon de noter que l'introduction de coûts aléatoires peut causer des problèmes. Si les propriétés de convexité des fonctions de Bellman ne sont pas conservées (ce qui est le cas si les prix sont corrélés entre eux), il n'est plus possible de faire une approximation de ces fonctions par

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

des plans coupants. Dans ce cas, soit le coût doit demeurer fixe soit nous appliquons des approches existantes combinant programmation dynamique et programmation dynamique duale.

Une autre éventualité est l'optimisation robuste. Le problème de gestion d'une vallée hydraulique par **optimisation robuste** a déjà été étudié au chapitre 4. Une telle approche permet d'obtenir une capacité maximale de production robuste à l'aléa d'apports. Avec l'introduction d'ajustements linéaires, ce problème est rendu réaliste et adaptée à la réalisation des scénarios d'apports. Nous avons d'ailleurs vu au chapitre 4 que cette modélisation donnait de meilleurs résultats en terme de satisfaction de contraintes qu'une approche déterministe ou même déterministe glissante. Le grand avantage de cette approche est qu'une unique optimisation est nécessaire (si les coûts sont fixes ou si les coûts sont également gérés par optimisation robuste). La reconstruction des commandes ne se fait que par simples simulations sur chaque scénario d'apports. Une telle approche peut donc s'avérer très intéressante et pourrait même réduire les temps de calcul importants du calcul de la capacité de production maximale hydraulique. Enfin, si nous considérons des scénarios de prix, il faudra reconstruire les commandes sur l'ensemble des scénarios d'apports pour chaque scénario de prix. Ce cas de figure peut par contre être très coûteux et donc nécessiter la réduction du nombre de scénarios pris en compte.

9.2.2 Boucle ouverte : calcul de marge intégré au problème de couverture

Dans le chapitre 5, la marge est calculée a priori : elle est une donnée du problème sous forme de scénarios. Or, cette façon de faire contraint le décideur à ne pouvoir utiliser que des actifs financiers de couverture, actifs qui pourraient s'avérer plus onéreux que des actifs physiques. Des recours possibles seraient la modification du planning de maintenance des centrales ou une modification de la politique de turbinage des usines hydrauliques. Mais pour cela, nous devons introduire le calcul de marge au sein même du problème de couverture. Nous étudions cette possibilité dans cette section.

Capacité maximale de production hydraulique

La capacité de production maximale hydraulique est complexe à calculer. Nous rappelons qu'il est nécessaire de résoudre le problème d'optimisation d'une vallée hydraulique à chaque jour, pour chaque vallée et pour un scénario d'apports donné. Intégrer ce modèle en l'état semble réellement difficile, puisque calculé au sein d'une contrainte en probabilité. Pour espérer atteindre cet objectif, il peut être nécessaire de trouver une autre approche de calcul de la capacité maximale hydraulique ou d'imaginer un problème à plusieurs niveaux. Il est assez clair que, comme l'hydraulique est disponible rapidement à moindre coût, les enjeux d'une telle modélisation pourraient être très importants. Mais pour l'heure cette idée n'est qu'une piste de recherche de grande complexité.

Capacité maximale de production thermique

La problématique de la capacité maximale de production thermique semble beaucoup plus accessible dans un premier temps. Dans la situation où le critère de risque appliqué n'est pas satisfait, il est possible de récupérer de la puissance disponible en déplaçant, dans certaines mesures, des arrêts de centrale prévus pour maintenance. En effet, nous pouvons, par ce biais, rajouter de la puissance à un instant où la marge est déficiente et cela pour un coût bien inférieur au coût d'achat de produits financiers.

Les centrales déplaçables sont :

- les centrales démarrées en t_0 et ayant un arrêt prévu sur la période d'étude ;
- les centrales non démarrées en t_0 mais ayant quand même un arrêt prévu sur cette même période.

Nous rappelons que la marge thermique est la somme des productions maximales de toutes les centrales disponibles au jour j .

Modèle Soit :

- Θ l'ensemble des centrales thermiques disponibles à la production n'ayant pas d'arrêt prévu pendant la période Ω_J ;
- Θ_{arr} l'ensemble des centrales thermiques disponibles ayant un arrêt prévu.

La centrale $i \in \Theta_{arr}$ a donc un arrêt prévu à la date t_i^A et de durée d_i . Nous pouvons déplacer cet arrêt dans l'intervalle $\mathcal{D} = [\max(1, t_i^A - \underline{\epsilon}_i); \min(J, t_i^A + \bar{\epsilon}_i)]$.

Pour savoir si une centrale est disponible ou non au jour j , nous introduisons la variable binaire b_i^j :

$$b_i^j = \begin{cases} 1 & \text{si } j \text{ est le début de l'arrêt de la centrale } i \\ 0 & \text{sinon.} \end{cases}$$

Pour savoir si la date j est dans la période de maintenance, il faut parcourir l'intervalle $[j - d_i + 1; j]$ et voir si la date de début de la maintenance y est incluse. La quantité $\mathbb{1}_i^j = 1 - \sum_{j'=j-d_i+1}^j b_i^{j'}$ joue donc le rôle de fonction indicatrice qui vaut :

$$\mathbb{1}_i^j = \begin{cases} 0 & \text{si } j \text{ est dans la période de maintenance de } i \\ 1 & \text{sinon.} \end{cases}$$

Une centrale peut d'autre part tomber en panne pendant la période d'étude. Introduire le modèle d'indisponibilité utilisé pour le calcul de la marge dans la contrainte en probabilité est très difficile. En effet, avec un tel modèle, l'indisponibilité deviendrait aléatoire et nous aurions un produit de variables avec des coefficients aléatoires. Dans un premier temps, nous proposons une alternative simple et relativement réaliste. Nous choisissons pour cela de faire un abattement de puissance $\alpha_i = 1 - \nu_i \cdot e^{-\nu_i}$ (suivant donc une loi de Poisson) avec un taux de panne moyen ν_i par centrale. Nous pouvons donc définir la marge thermique C_{\max}^{Th} au jour j :

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

$$C_{\max,j}^{Th} = \sum_{i=1}^{|\Theta|} \alpha_i P_{ij}^{\max} + \sum_{i=1}^{|\Theta_{arr}|} \alpha_i P_{ij}^{\max} \mathbb{1}_i^j, \quad \forall j \in \Omega_J.$$

Pour déplacer un arrêt, l'optimiseur doit attribuer la valeur 1 à la variable b_{ij} correspondant à la nouvelle date de début d'arrêt j (forcément située dans l'intervalle \mathcal{D}) et à 0 partout ailleurs (il ne peut y avoir qu'un unique arrêt sur la période). Il nous faut donc rajouter ces contraintes de limitation de déplacements :

$$\begin{cases} \sum_{j \in \mathcal{D}} b_i^j = 1, \quad \forall i \in \Theta_{arr} \\ b_i^j = 0, \quad \forall i \in \Theta_{arr}, \quad \forall j \in \Omega_J \setminus \mathcal{D} \end{cases}$$

Nous pouvons également limiter le nombre de déplacements de centrales sur la période d'étude à un nombre N^A fixé par avance. Pour cela, il suffit de sommer le nombre de dates de début d'arrêt b_i^j qui soient différents de la date t_i^A prévue :

$$\sum_{i \in \Theta_{arr}} \sum_{\substack{j \in \mathcal{D} \\ j \neq t_i^A}} b_i^j \leq N^A.$$

Enfin, même si ces déplacements ont un coût négligeable pour le producteur, le nombre de déplacements autorisés est très limité et il nous faut donc pouvoir arbitrer économiquement entre déplacer ou non. Pour cela, nous facturons la production thermique journalière à un certain coût ρ_i^j (pouvant être par exemple le coût marginal du système). Le coût de la production thermique sur la période Ω_J est donc :

$$c^\theta = \sum_{i \in \Theta_{arr}} \sum_{j \in \Omega_J} \alpha_i \rho_i^j P_{ij}^{\max} \mathbb{1}_i^j.$$

Cependant, c^θ représente le coût de production thermique journalier. Ce coût permet d'effectuer les divers arbitrages mais n'est pas représentatif. Ici, seul le coût de déplacement, et non de production, nous intéresse puisque nous cherchons à minimiser le coût de couverture. Le coût de déplacement est représenté par l'écart entre le coût du planning original et le coût du planning modifié. Le véritable coût s'obtient donc en rajoutant le terme constant représentant le coût de planning de maintenance initial.

Soit ψ_i^j définit par :

$$\psi_i^j = \begin{cases} 0 & \text{si } j \in [t_i^A; t_i^A + d_i - 1] \\ 1 & \text{sinon.} \end{cases}$$

Le coût thermique s'écrit donc :

$$c^\theta = \sum_{i \in \Theta_{arr}} \sum_{j \in \Omega_J} \alpha_i \rho_i^j (P_{ij}^{\max} \mathbb{1}_i^j - P_{ij}^{\max} \psi_i^j).$$

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

Comme les pannes sont équivalentes à un abattement fixe de puissance, la seule variable aléatoire dans la contrainte en probabilité est la marge tronquée $\hat{M}_t = C_{\max}^{Hyd}(\tilde{\xi}^I, \tilde{\xi}^H) - \tilde{D}_t$, qui est donc la marge sans la capacité maximale de production thermique. Nous nous retrouvons en fait dans le même cas que dans le cas général (variable aléatoire séparable, scalaire) et nous pouvons donc utiliser une approche par quantile (voir section 5-5.4.2). Nous pouvons donc récapituler le nouveau modèle :

$$\begin{aligned}
 \min_{q^A, q^V, e, b} \quad & \sum_{j=1}^{J-1} \sum_{c \in \mathcal{C}^S} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + \sum_{j=1}^{J-1} \sum_{k=1}^{|OT|} p_{k,j}^{OT} q_{k,j}^{OT} e_j^k + 5 \sum_{j \in \Omega_J^-} \sum_{c \in \mathcal{C}^W} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) \\
 & + 2 \sum_{j \in \Omega_J^{We}} \sum_{c \in \mathcal{C}^{We}} (\bar{p}_j^{A,c} q_j^{A,c} - \bar{p}_j^{V,c} q_j^{V,c}) + \sum_{i \in \Theta_{arr}} \sum_{j \in \Omega_J} \alpha_i \rho_i^j (P_{ij}^{\max} \mathbb{1}_i^j - P_{ij}^{\max} \psi_i^j). \\
 \text{s.t.} \quad & \left\{ \begin{aligned} & q_M^j(p) + \sum_{k \in S_j^-} \sum_{c \in \mathcal{C}^W} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \notin \{\text{Sa}, \text{Di}\}\}} + \sum_{k \in S_j} \sum_{c \in \mathcal{C}^{We}} (q_k^{A,c} - q_k^{V,c}) \mathbb{1}_{\{j \in \{\text{Sa}, \text{Di}\}\}} \\ & + \sum_{c \in \mathcal{C}^S} (q_{j-1}^{A,c} - q_{j-1}^{V,c}) + \sum_{k=1}^{|OT|} q_{k,j-1}^{OT} e_{j-1}^k + \sum_{i=1}^{|\Theta|} \alpha_i P_{ij}^{\max} + \sum_{i=1}^{|\Theta_{arr}|} \alpha_i P_{ij}^{\max} \mathbb{1}_i^j \geq M_S, \\ & \sum_{j \in \mathcal{D}} b_i^j = 1, \quad \forall i \in \Theta_{arr} \\ & b_i^j = 0, \quad \forall i \in \Theta_{arr}, \forall j \in \Omega_J \setminus \mathcal{D} \\ & \sum_{i \in \Theta_{arr}} \sum_{\substack{j \in \mathcal{D} \\ j \neq t_i^A}} b_i^j \leq N^A \\ & \sum_{j=1}^{J-1} e_j^k = 1, \quad \forall k = 1, \dots, |OT| \\ & \sum_{k=1}^{|OT|} \sum_{j=1}^{J-1} e_j^k \leq N_{OT} \\ & \forall c \in \mathcal{C}^S, \quad q^{A,c}, q^{V,c} \in U_S \\ & \forall c \in \mathcal{C}^W, \quad q^{A,c}, q^{V,c} \in U_W \\ & \forall c \in \mathcal{C}^{We}, \quad q^{A,c}, q^{V,c} \in U_{We} \end{aligned} \right. \quad \forall j \in \Omega_J
 \end{aligned} \tag{9.1}$$

Résultats Nous construisons un jeu d'essais sur 12 jours qui permet de bénéficier d'une marge initiale équivalente à la section (5.4.4) du chapitre 5. Nous conservons un marché Spot (Base et Pointe), 1 contrat Week-Ahead Base et 1 contrat Week-Ahead Pointe, 1 contrat Week-end Ahead Base, et 10 options tarifaires. Nous introduisons une puissance fixe thermique de 1500MW (des centrales disponibles sans arrêt prévu) et de 6 centrales thermiques dont un arrêt est programmé :

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

Centrale	Date de début d'arrêt	Intervalle de déplacement	Durée de l'arrêt	P_{\max}	Coût par jour de production
1	3	[2 ; 7]	3 jours	800 MW	90 euros\MW
2	5	[4 ; 9]	3 jours	400 MW	70 euros\MW
3	7	[6 ; 11]	4 jours	300 MW	50 euros\MW
4	7	[6 ; 11]	3 jours	500 MW	80 euros\MW
5	9	[8 ; 12]	5 jours	400 MW	70 euros\MW
6	5	[3 ; 9]	4 jours	500 MW	80 euros\MW

TAB. 9.1 – Description du parc thermique disponible avec des arrêts pour maintenance prévus

Avec aucun déplacement d'arrêt, nous retrouvons les commandes de couverture proposées en (5.4.4) (Fig. 9.2) :

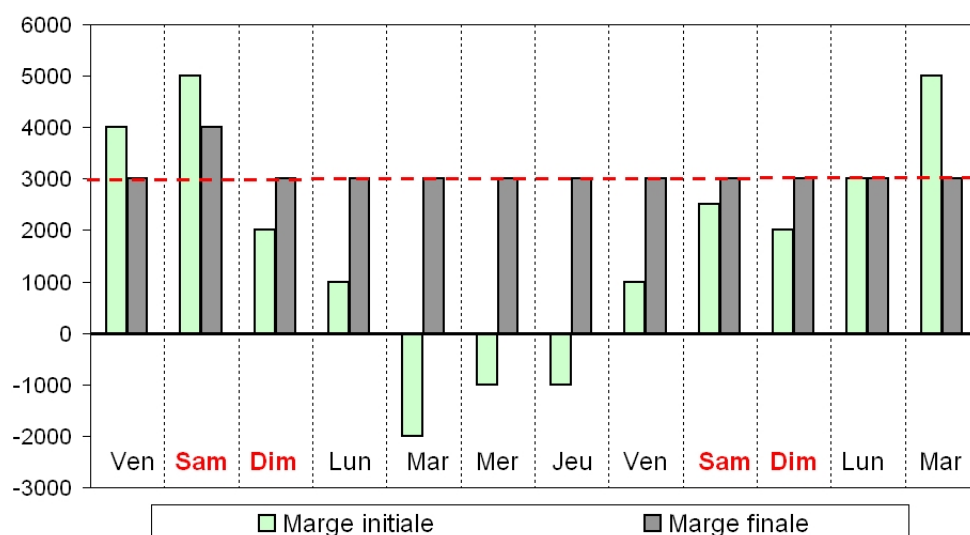


FIG. 9.1 – Comparaison de la marge de production entre avant et après l'application des décisions de couverture

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

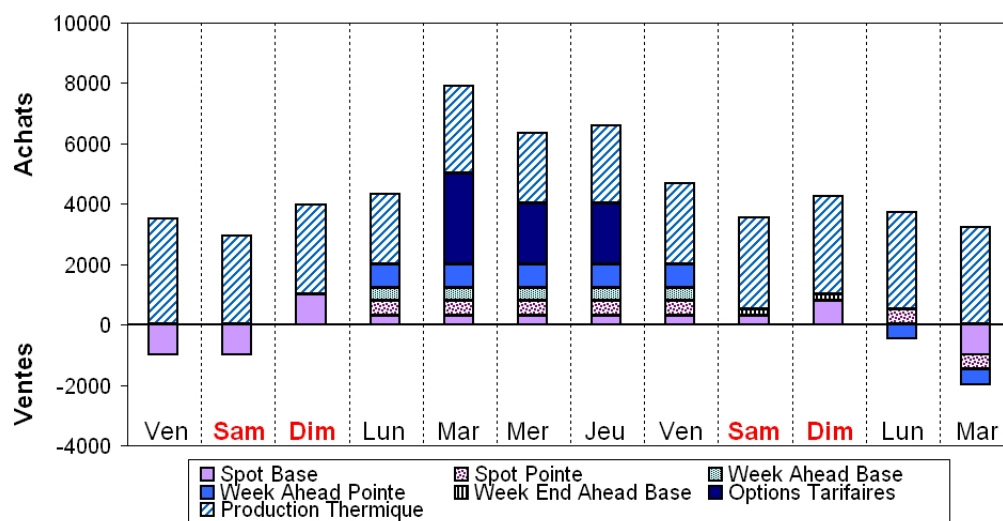


FIG. 9.2 – Description des commandes de couverture appliquées (aucun déplacement d'arrêts de centrale autorisé)

Le coût de couverture est de 21 865 593 euros et le critère de marge est bien, au final, satisfait (Fig. 9.1). Désormais, si nous autorisons, par exemple 4 déplacements d'arrêts pour maintenance de centrales, les résultats obtenus sont les suivants :

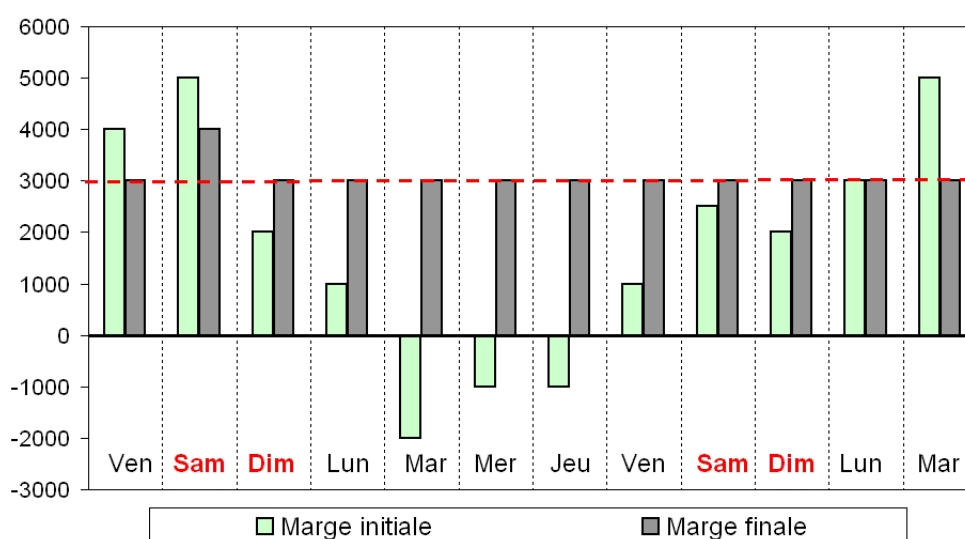


FIG. 9.3 – Comparaison de la marge de production entre avant et après l'application des décisions de couverture (4 déplacements d'arrêts autorisés)

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

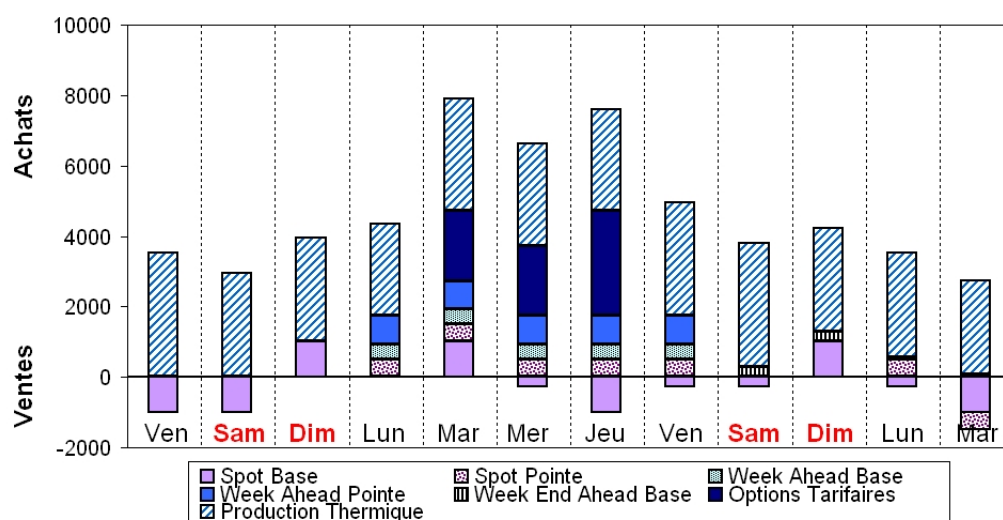


FIG. 9.4 – Description des commandes de couverture appliquées (quatre déplacements d’arrêts de centrale autorisés)

Même si le niveau de marge final est équivalent (Fig. 9.3) à la situation sans arrêt autorisé, les centrales 2 à 6 ont été déplacées, contraignant l’optimiseur à trouver de nouvelles commandes financières (Fig. 9.4). Les opportunités de vente sont plus nombreuses puisque le modèle déplace des arrêts pour récupérer de la puissance sur des jours où le prix de vente sur les marchés est plus élevé. Nous faisons ainsi des gains non négligeables puisque le coût final, pour un même niveau de couverture que précédemment, est de 16 381 334 euros soit une réduction des coûts de 25%. Cet exemple simple illustre l’intérêt de prendre en compte ces déplacements d’arrêts. Dans le tableau 9.2 suivant, nous récapitulons les gains occasionnés en fonction du nombre de déplacements sachant que la marge finale demeure toujours supérieure aux 3000 MW réglementaires. Les réductions de coûts peuvent aller jusqu’à 27% :

Nombre de déplacements autorisés	Coût final de la couverture	Ecart de coût avec un parc figé
0	21 865 593 euros	0%
1	18 810 548 euros	14%
2	17 631 176 euros	19%
3	16 900 242 euros	23%
4	16 381 334 euros	25%
5	15 978 160 euros	27%
6	15 978 160 euros	27%

TAB. 9.2 – Coûts de la couverture en fonction du nombre de déplacements d’arrêts autorisés

Nous montrons donc ici qu’il existe un réel intérêt à prendre en compte le calcul de marge à l’intérieur du modèle d’optimisation. Le bémol à poser sur cette modélisation est la prise en

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

compte de l'aléa d'indisponibilité. Pour simplifier les choses, nous avons fait un abattement de puissance constant. Pour améliorer le modèle, il est nécessaire d'intégrer un véritable modèle aléatoire d'indisponibilité. Cependant, avec la contrainte en probabilité, nous risquons de nous heurter à de nombreuses difficultés.

9.2.3 Boucle fermée

La partie sur la formulation du problème en boucle fermée n'est qu'une première piste d'exploration. Nous avons proposé une approche qui fournit des résultats intéressants mais, finalement, tout reste à faire dans cette partie, notamment en comparant diverses approches. Voici quelques propositions qui nous sembleraient intéressantes.

Programmation stochastique avec règles de décision constantes par morceaux

La première piste est bien entendu de poursuivre les expérimentations du chapitre 6. Les outils étant mis en place, il faut continuer à multiplier les expériences numériques en jouant sur les différents paramètres (nombre et taille des experts, nombre d'ensembles de décision, ...). Nous pouvons aussi affiner la modélisation du problème en considérant plus de contrats voire même en essayant d'intégrer des variables entières pour rendre plus réaliste la gestion des options tarifaires.

L'optimisation robuste

Nous avons également introduit au chapitre 6 un premier modèle robuste simple. Cette approche permet de satisfaire les niveaux de contrainte tout en dégagant des bénéfices. Cependant, ce modèle est sur-protecteur et il serait intéressant d'intégrer de l'adaptation avec, notamment, des règles de décision.

A l'image des travaux sur l'hydraulique (chapitre 4), nous introduisons des variables de décision qui sont dépendantes linéairement des réalisations passées afin d'utiliser l'information disponible à un instant t . Si nous faisons ainsi dépendre les commandes des s dernières réalisations de l'aléa de marge, on peut obtenir le problème robuste (9.3) :

$$\left\{ \begin{array}{l} x_t = x_t^0 + \sum_{\substack{\tau=1 \\ t-\tau>0}}^s x_t^\tau \epsilon_{t-\tau}, \quad \forall t \in 1, \dots, T-1 \\ y_t = y_t^0 + \sum_{\substack{\tau=1 \\ t-\tau>0}}^s y_t^\tau \epsilon_{t-\tau}, \quad \forall t \in 1, \dots, T/2 \\ z_t = z_t^0 + \sum_{\substack{\tau=1 \\ t-\tau>0}}^s z_t^\tau \epsilon_{t-\tau}, \quad \forall t \in 1, \dots, T-1. \end{array} \right. \quad (9.2)$$

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

$$\begin{aligned}
& \bar{\eta}_{\tau,t}^{xp} + \bar{\zeta}_t^x \geq x_t^{t-\tau}, \quad t = 1, \dots, T-1, \quad \tau = \max(t-s, 1), \dots, t-1 \\
& \bar{\eta}_{\tau,t}^{xm} + \bar{\zeta}_t^x \geq -x_t^{t-\tau}, \quad t = 1, \dots, T-1, \quad \tau = \max(t-s, 1), \dots, t-1 \\
& x_t^0 + \kappa \zeta_t^x \sqrt{t - \max(t-s, 1)} + \kappa \sum_{\tau=\max(t-s, 1)}^{t-1} (\eta_{\tau,t}^{xp} + \eta_{\tau,t}^{xm}) \geq \underline{x}_t, \quad \forall t = 1, \dots, T-1 \\
& \underline{\eta}_{\tau,t}^{xp} + \underline{\zeta}_t^x \geq x_t^{t-\tau}, \quad t = 1, \dots, T-1, \quad \tau = \max(t-s, 1), \dots, t-1 \\
& \underline{\eta}_{\tau,t}^{xm} + \underline{\zeta}_t^x \geq -x_t^{t-\tau}, \quad t = 1, \dots, T-1, \quad \tau = \max(t-s, 1), \dots, t-1 \\
& u_{\tau,t}^p, u_{\tau,t}^m \geq 0, \quad v_t \geq 0, \quad \rho \geq 0, \quad \lambda_t^p, \lambda_t^m \geq 0, \quad \forall t = 1, \dots, T-1, \quad \forall \tau = 1, \dots, t,
\end{aligned} \tag{9.4}$$

avec x^\bullet le vecteur constitué de toutes les variables x_t^τ définies dans les règles (de même pour y^\bullet et z^\bullet).

Appliquons cette approche sur le jeu d'essais présenté précédemment (cf. section 6-6.6.1). Voici un tableau récapitulatif de résultats :

Nombre de termes dans la règle	Couverture	Pourcentages de violations de la contrainte en proba.								Coût moyen
1,3,5	99%	0,84%	0%	0%	0,46%	0%	0%	0%	0%	1 857 648€

TAB. 9.3 – Niveau de violations de la contrainte du risque 1% et coût moyen sur 75 000 scénarios pour une approche RAL

Plusieurs constats à la vue de ce tableau :

- le risque 1% est quasi-vérifié avec une couverture à 99%, ce qui est satisfaisant ;
- un pas de temps d'adaptation dans la règle suffit puisque plus de connaissances ne change plus les résultats ;
- le nombre global de violations est particulièrement réduit puisqu'elles sont nulles sur 7 pas de temps ;
- le coût de la couverture est désastreux : au lieu de dégager de bénéfices, désormais nous devons payer, et même très cher.

Il est évident que ce modèle, bien que satisfaisant en terme de violations, est trop conservatif et conduit à une mauvaise gestion économique. Pour mieux comprendre cet écart de coût, nous pouvons comparer les décisions du modèle robuste de la section 6-6.6.2 avec celles de ce modèle RAL sur un scénario quelconque de marge :

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

t	1	2	3	4	5	6	7	8	9
Marge (MW)	7363.14	15373.44	18670.8	9191.9	7670.56	5713.76	4619.142	7129.58	10591.24
Robuste SPOT	-2000	-2000	-2000	-2000	-123.74	1405.76	2961.3	3683.04	495
RAL SPOT	-1855.80	-2000	-2000	524.35	-366.34	1374.72	4038.5	6000	4310.2
Robuste WA	-150	-150	-150	-150	-150				
RAL WA	500	500	500	500	474.8				
Robuste OT	0	0	0	0	0	0	0	0	0
RAL OT	0	0	0	0	0	0	0	500	0

TAB. 9.4 – Commandes de SPOT, WA et OT avec des modèles robuste et RAL

Bien qu’adaptatif, le modèle RAL achète excessivement sur un scénario où, à la base, la marge est excédentaire. Nous pouvons essayer d’améliorer cela en considérant, non plus une couverture globale mais plutôt une couverture pas de temps par pas de temps : nous conservons une couverture à 99% les jours où les violations initiales sont importantes et nous la réduisons pour les autres en fonction de la marge initiale. À l’image du modèle robuste 6-6.6.2, nous pouvons obtenir de meilleurs résultats en modifiant les niveaux de couverture par tâtonnement. Cette approche peut permettre de dégager des bénéfices mais est peu réaliste en pratique.

En fait, nous pouvons expliquer assez simplement pourquoi les modèles RAL sont peu satisfaisants dans ce cas. Dans les modèles robustes appliqués à l’hydraulique, les contraintes à rendre robustes portaient sur des états faisant intervenir toute l’historique des aléas. Donc l’intégration d’aléas via les règles de décision n’augmentaient en rien le nombre d’aléas et enrichissait la prise de décision. Par contre, dans le cas de la marge, c’est tout à fait différent. La marge ne dépend que d’un seul instant, ce qui signifie que si la règle dépend d’aléas supplémentaires, nous rajoutons artificiellement de l’incertitude au problème. Ce qui a pour effet d’accroître la taille de l’ensemble d’incertitude et donc trouver des solutions plus conservatives. D’autre part, la règle apporte une incertitude contraignant encore plus le problème. Mais comme cet aléa est multiplicatif de variables, l’optimiseur peut facilement la supprimer en annulant tous les coefficients de la règle. Tout l’intérêt d’intégrer des règles de décision disparaît donc.

Il semble donc que des règles de décision telles que nous les avons utilisées dans le chapitre sur l’hydraulique (chapitre 4) ne soient pas forcément adaptées à ce cas, même si les critères de risque sont bien satisfaits. Nous pouvons voir deux façons de résoudre ce problème :

- utiliser de nouveaux types de règles : des règles du style $x_t = x^0 + \epsilon_{t-1}$ résoudraient ce problème. Par contre, nous pouvons nous demander si elles sont censées et si elles aident à la résolution. Seule l’expérimentation peut le montrer ; sinon d’autres types peuvent être imaginées ;
- modéliser la marge par des modèles plus riches, style autorégressifs. Nous introduisons ainsi, à l’image d’un réservoir hydraulique, toute l’historique de l’aléa. Nous avons déjà discuté au chapitre 7 de la difficulté de caler de tels modèles. Mais ce constat montre qu’il y aurait vraiment un intérêt à faire une étude statistique poussée du processus de marge.

Ces premiers modèles robustes simples sont imparfaits mais indique bien qu’il y a un potentiel à utiliser des méthodes robustes.

Couplage de l'optimisation robuste et de la programmation stochastique

Pour gérer au mieux l'aléa de prix, nous intégrons une règle de décision sur le prix. Mais pour ne pas retomber dans les travers de l'approche précédente, nous gérons cet aléa par programmation stochastique. Les prix p sont donc représentés par un peigne de scénarios.

Pour simplifier, nous ne supposons qu'une seule commande d'achats/ventes de spot x_t pour livraison en $t + 1$. Nous introduisons la règle de décision :

$$x_t = x_t^0 + x_t^1 p_{t-1}.$$

Le problème initial s'écrit sous une forme simplifiée :

$$\begin{aligned} \min_{x,y,z} \quad & \mathbb{E}\left[\sum_{t=1}^{T-1} p_t x_t\right] \\ \text{s.t.} \quad & \begin{cases} \mathbb{P}[M_t + x_{t-1} \geq \mathcal{M}_S] \geq 1 - p, \quad \forall t \in 2, \dots, T \\ \underline{x}_t \leq x_t \leq \bar{x}_t, \quad \forall t = 1, \dots, T-1. \end{cases} \end{aligned} \tag{9.5}$$

En introduisant la règle de décision linéaire, le problème devient :

$$\begin{aligned} \min_{x,y,z} \quad & \mathbb{E}\left[\sum_{t=1}^{T-1} p_t (x_t^0 + x_t^1 p_{t-1})\right] \\ \text{s.t.} \quad & \begin{cases} M_t + \sigma_t \epsilon_t + x_{t-1}^0 + x_{t-1}^1 p_{t-2} \geq \mathcal{M}_S, \quad \forall \epsilon_t \in \Xi_t, \quad \forall t \in 2, \dots, T \\ \underline{x}_t \leq x_t^0 + x_t^1 p_{t-1} \leq \bar{x}_t, \quad \forall t = 1, \dots, T-1, \end{cases} \end{aligned} \tag{9.6}$$

où Ξ_t est un ensemble d'incertitude calibré de sorte à approximer la contrainte en probabilité précédente.

Une approximation de l'objectif peut être :

$$\mathbb{E}\left[\sum_{t=1}^{T-1} p_t (x_t^0 + x_t^1 p_{t-1})\right] = \frac{1}{N} \left[\sum_{t=1}^{T-1} \sum_{n=1}^N p_t^n (x_t^0 + x_t^1 p_{t-1}^n) \right].$$

La contrainte de marge est rendue robuste par rapport à l'aléa de marge alors que les prix sont gérés par programmation stochastique. Nous avons donc autant de contraintes de marge que de scénarios.

La contrainte de marge devient :

$$M_t - \kappa_t \sigma_t + x_{t-1}^0 + x_{t-1}^1 p_{t-2}^n \geq \mathcal{M}_S, \quad \forall n = 1, \dots, N, \quad \forall t \in 2, \dots, T.$$

Enfin, les contraintes de bornes deviennent :

9.2. PERSPECTIVES SUR LA GESTION DES RISQUES DE DÉFAILLANCE PHYSIQUE

$$\underline{x}_t \leq x_t^0 + x_t^1 p_{t-1}^n \leq \bar{x}_t, \quad \forall t = 1, \dots, T-1, \quad \forall n = 1, \dots, N.$$

Cette approche mériterait d'être testée et devrait permettre de gérer conjointement les deux aléas sans rendre le modèle trop contraint.

Cinquième partie

Annexes

Annexe A

Application de l'optimisation robuste à la gestion hydraulique hebdomadaire : compléments

A.1 Equivalent robuste de l'objectif

Considérons la fonction objectif

$$J = \sum_r \omega_r (V_r^0 - V_r^H) + \sum_g \rho_g \sum_{\tau \leq H} c_\tau \tilde{F}_{g\tau}.$$

Explicitons préalablement les termes $V_r^0 - V_r^H$ et $\sum_g \rho_g \sum_{\tau \leq H} c_\tau \tilde{F}_{g\tau}$. Nous avons

$$\begin{aligned} V_r^0 - V_r^H &= - \left[\sum_{\tau \leq H} a_r^\tau + \sum_{j \leq J(H)+1} \sigma_r \sqrt{T} \theta_r^j + \sum_{\tau \leq H} \sum_g \delta_{gr} \tilde{F}_{g\tau} \right] \\ &= \left[\sum_{\tau \leq H} a_r^\tau + \sum_{j \leq J(H)+1} \sigma_r \sqrt{T} \theta_r^j + \right. \\ &\quad \left. + \sum_{\tau \leq H} \sum_g \delta_{gr} \left(F_{g\tau} + \sum_{l=1}^k x_{g\tau}^l \theta^{J(\tau)-l+1} \right) \right], \end{aligned}$$

et

$$\begin{aligned} \sum_g \rho_g \sum_{\tau \leq H} c_\tau \tilde{F}_{g\tau} &= \sum_g \rho_g \sum_{\tau \leq H} c_\tau \left[F_{g\tau} + \sum_{l=1}^k x_{g\tau}^l \theta^{J(\tau)-l+1} \right] \\ &= \sum_g \rho_g \sum_{\tau \leq H} c_\tau F_{g\tau} + \sum_g \rho_g \sum_{\tau \leq H} c_\tau \sum_{l=1}^k x_{g\tau}^l \theta^{J(\tau)-l+1} \end{aligned}$$

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION
HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS

Nous nous proposons maintenant de regrouper les termes indépendants des aléas dans la somme J , puis les termes dépendants des aléas. Cela donne

$$J = J_{det} + J_{sto}$$

avec

$$\begin{aligned} J_{det} &= - \sum_r \omega_r \left[\sum_{\tau \leq H} a_r^\tau + \sum_{\tau \leq H} \sum_g \delta_{gr} F_{g\tau} \right] + \sum_g \rho_g \sum_{\tau \leq H} c_\tau F_{g\tau} \\ &= \sum_{\tau \leq H} \left[- \sum_r \omega_r (a_r^\tau + \sum_g \delta_{gr} F_{g\tau}) + \sum_g \rho_g c_\tau F_{g\tau} \right] \end{aligned}$$

et

$$\begin{aligned} J_{sto} &= - \sum_r \omega_r \left[\sum_{j \leq J(H)+1} \sigma_r \sqrt{T} \theta_r^j + \sum_{\tau \leq H} \sum_g \delta_{gr} \sum_{l=1}^k x_{g\tau}^l \theta^{J(\tau)-l+1} \right] \\ &\quad + \sum_g \rho_g \sum_{\tau \leq H} c_\tau \sum_{l=1}^k x_{g\tau}^l \theta^{J(\tau)-l+1}. \end{aligned}$$

En utilisant la notation pour $\forall r, r' = 1 \dots R, \forall j = 1, \dots, J(H)$

$$\alpha_{r'r}^{Hj} = I_{r'r} \sigma_r \sqrt{T} + \sum_g \delta_{gr} \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} x_{g\tau}^l \right)$$

avec

$$\begin{aligned} s(l, j, H) &= \min\{(j+l)T, H\}, \\ I_{r'r} &= \begin{cases} 1 & \text{si } r' = r, \\ 0 & \text{autrement.} \end{cases} \end{aligned}$$

Nous pouvons réécrire

$$\begin{aligned} J_{sto} &= - \sum_r \omega_r \left[\sigma_r \sqrt{T} \theta_r^{J(H)+1} + \sum_{j \leq J(H)} \sum_{r'} \alpha_{r'r}^{Hj} \theta_{r'}^j \right] \\ &\quad + \sum_{j \leq J(H)} \sum_{r'} \left[\sum_g \rho_g \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} c_\tau x_{g\tau}^l \right) \right] \theta_{r'}^j \\ &= - \sum_r \omega_r \sigma_r \sqrt{T} \theta_r^{J(H)+1} + \sum_{j \leq J(H)} \sum_{r'} \left[- \sum_r \omega_r \alpha_{r'r}^{Hj} + \right. \\ &\quad \left. + \sum_g \rho_g \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} c_\tau x_{g\tau}^l \right) \right] \theta_{r'}^j \end{aligned}$$

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION
HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS

Nous introduisons l'expression

$$\beta_{r'}^j = - \sum_r \omega_r \alpha_{r'r}^{Hj} + \sum_g \rho_g \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} c_\tau x_{g\tau}^l \right).$$

Cela permet d'écrire la condition de profit minimal garanti sous la forme compacte

$$z \geq J_{det} + \sum_{j \leq J(H)} \sum_{r'} \beta_{r'}^j \theta_{r'}^j.$$

L'équivalent robuste s'obtient en résolvant le problème

$$\max \left\{ \sum_{r'=1}^R \sum_{j \leq J(H)} \beta_{r'}^j \theta_{r'}^j \mid \sum_{j=1}^{J(H)} \sum_{r'=1}^R |\theta_{r'}^j| \leq \kappa \sqrt{J(H)R}, |\theta_{r'}^j| \leq \kappa, \forall j \in 1, \dots, J(H), \forall r' \in 1, \dots, R \right\}.$$

En suivant le même raisonnement que pour les contraintes précédentes, l'équivalent robuste de la contrainte de valeur garantie s'écrit :

$$J_{det} + v \kappa \sqrt{J(H)R} + \kappa \sum_{r'=1}^R \sum_{j \leq J(H)} (u_{r',j})_p + \kappa \sum_{r'=1}^R \sum_{j \leq J(H)} (u_{r',j})_m \leq z$$

$$\begin{aligned} (u_{r',j})_p + v &\geq \beta_{r'}^j, \forall j \in 1, \dots, J(H), \forall r' \in 1, \dots, R \\ (u_{r',j})_m + v &\geq -\beta_{r'}^j, \forall j \in 1, \dots, J(H), \forall r' \in 1, \dots, R \\ (u_{r',j})_p &\geq 0, (u_{r',j})_m \geq 0, v \geq 0 \end{aligned}$$

avec

$$J_{det} = \sum_{\tau \leq H} \left[- \sum_r \omega_r (a_r^\tau + \sum_g \delta_{gr} F_{g\tau}) + \sum_g \rho_g c_\tau F_{g\tau} \right]$$

et pour $\forall r, r' = 1 \dots R, \forall j = 1, \dots, J(H)$

$$\alpha_{r'r}^{Hj} = I_{r'r} \sigma_r \sqrt{T} + \sum_g \delta_{gr} \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} x_{g\tau}^l \right)$$

et pour $\forall r' = 1 \dots R, \forall j = 1, \dots, J(H)$

$$\beta_{r'}^j = - \sum_r \omega_r \alpha_{r'r}^{Hj} + \sum_g \rho_g \left(\sum_{l=1}^{\min(j,k)} \sum_{\tau=(j+l-1)T+1}^{s(l,j,H)} c_\tau x_{g\tau}^l \right).$$

Dans ces expressions, nous avons

$$\begin{aligned} s(l, j, H) &= \min\{(j+l)T, H\}, \\ I_{r'r} &= \begin{cases} 1 & \text{si } r' = r, \\ 0 & \text{autrement.} \end{cases} \end{aligned}$$

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS

Remarque : nous pouvons justifier ici l'hypothèse c_τ fixe. Si nous remplaçons dans l'objectif la commande par une règle de décision dépendant de l'aléa θ , nous obtenons un terme en $\theta.c_\tau$, produit de deux aléas, rendant l'optimisation impossible.

A.2 Etude du comportement du modèle robuste (type RAL) selon la couverture de l'objectif

Rendre robustes les contraintes de volume et de turbinage paraît logique puisqu'elles sont sujettes à des limites techniques. Par contre, il est plus difficile de justifier l'intérêt de la robustesse de la contrainte d'objectif. Nous allons donc fixer le coefficient de robustesse à 90% pour toutes les contraintes et faire varier celui de l'objectif pour étudier son influence.

En termes de violations de contraintes sur les contraintes de volume, le tableau Tab.A.1 montre bien que la variation de ce coefficient sur l'objectif est peu influent :

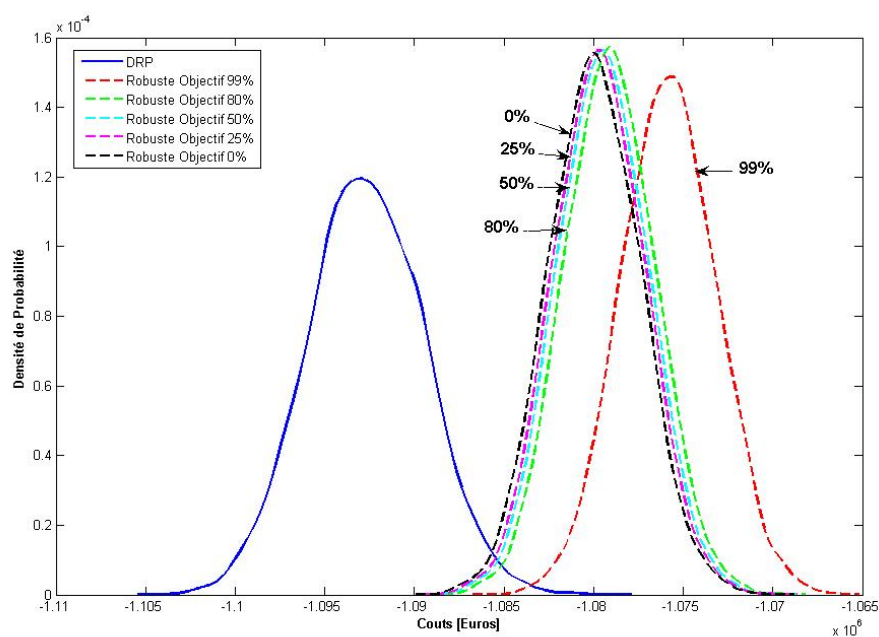
Contraintes violées	Taux de couverture de l'objectif					
	99%	95%	80%	50%	25%	0 %
Déversés [$m^3/scen$]	8,53	8,52	10,23	9,57	9,33	10,41
Violations V_{min} [$m^3/scen$]	0,395	0,395	0,395	0,395	0,395	0,395
Coût Moyen [Euros]	-1 075 811	-1 075 811	-1 079 094	-1 079 476	-1 079 686	-1 079 953

TAB. A.1 – Variation du coefficient de robustesse de l'objectif avec des contraintes robustes à 90%

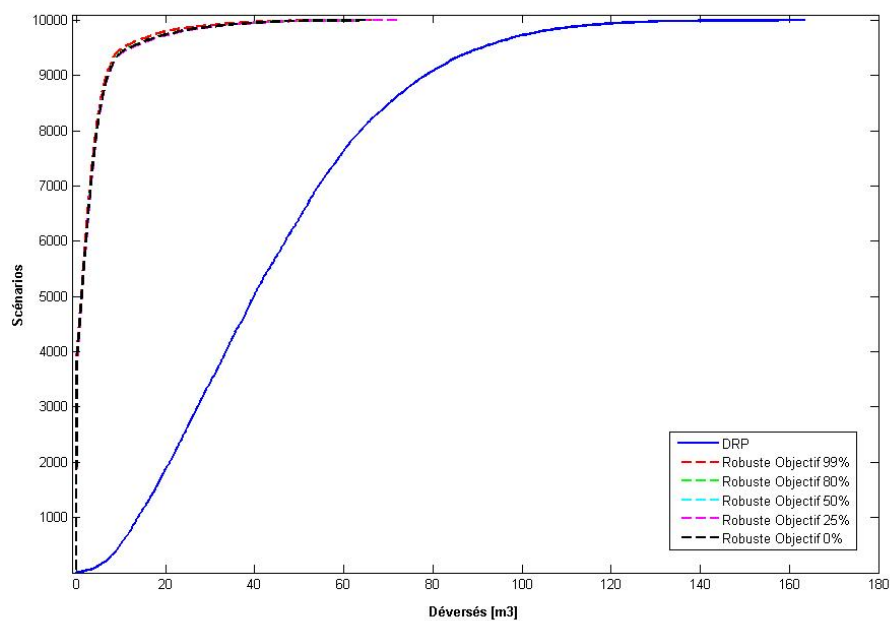
Si nous comparons les extrêmes (99% et 0%), ne pas rendre robustes les contraintes accroît sensiblement les gains pour une différence de déversés très faible. En effet, les déversés sont un peu plus importants puisque, ne rendant pas robuste la contrainte d'objectif, le modèle trouve une solution plus avantageuse financièrement mais aussi plus proche des bornes sur les contraintes (comme le fait le modèle déterministe) s'exposant donc plus aux violations. Mais nous pouvons voir ici que les conséquences de cette exposition sont très faibles ($2m^3$ en moyenne...). Nous pouvons donc raisonnablement économiser des variables avec un objectif non robuste.

Les figures suivantes confortent cette conclusion :

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS



(a)



(b)

FIG. A.1 – Comparatif des distributions de coûts (a) et des fonctions de répartition des déversés cumulé par scénario (b) du modèle robuste 90% avec variation de la couverture sur l'objectif.

A.3 Etude du comportement du modèle robuste (type RAL) selon le nombre de termes dans la règle de décision

Nous voulons étudier l'influence du nombre de termes dans la règle de décision sur les résultats. Nous n'introduisons pas de robustesse sur l'objectif et nous comparons différents indicateurs pour une couverture de 90% et une couverture de 50% sur toutes les contraintes.

Deux constats :

- le gain en terme de violations de contrainte n'est pas significatif :

Contraintes violées	K=1	K=2	K=3	K=4	K=5
Déversés (couv. 90%) [$m^3/scen$]	10,41	9,93	9,93	9,93	9,93
V_{min} (couv. 90%) [$m^3/scen$]	0,395	0,395	0,395	0,395	0,395
Déversés (couv. 50%) [$m^3/scen$]	21,48	23,93	23,93	23,93	23,93
V_{min} (couv. 50%) [$m^3/scen$]	0,395	0,395	0,395	0,395	0,395

TAB. A.2 – Violations de contraintes comme fonction du nombre de termes dans la règle de décision linéaire

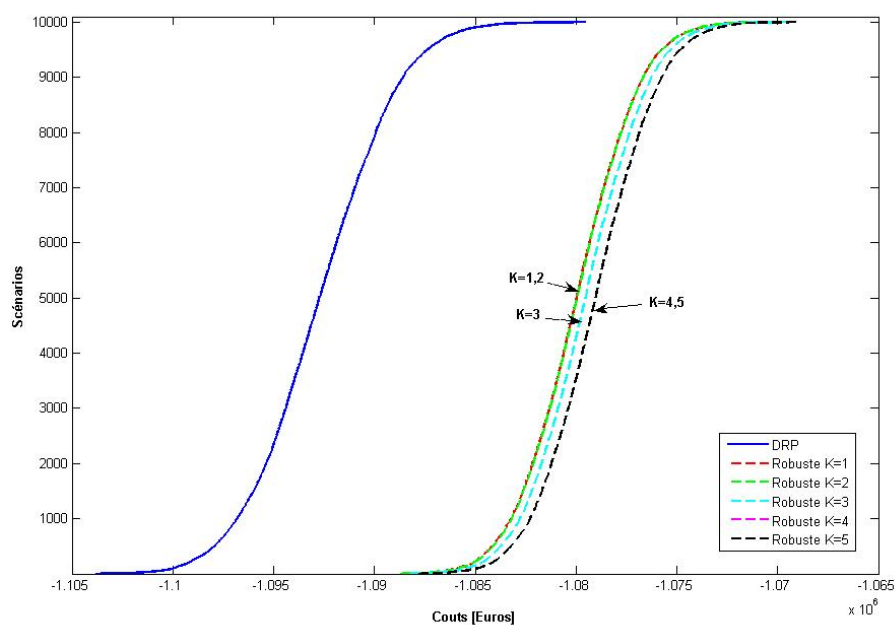
- La différence de coût est faible :

	K=1	K=2	K=3	K=4	K=5
Coût Moyen 90% [euros]	-1 079 952	-1 079 934	-1 079 518	-1 079 021	-1 079 021
Coût Moyen 50% [euros]	-1 083 044	-1 083 033	-1 082 796	-1 082 355	-1 082 355

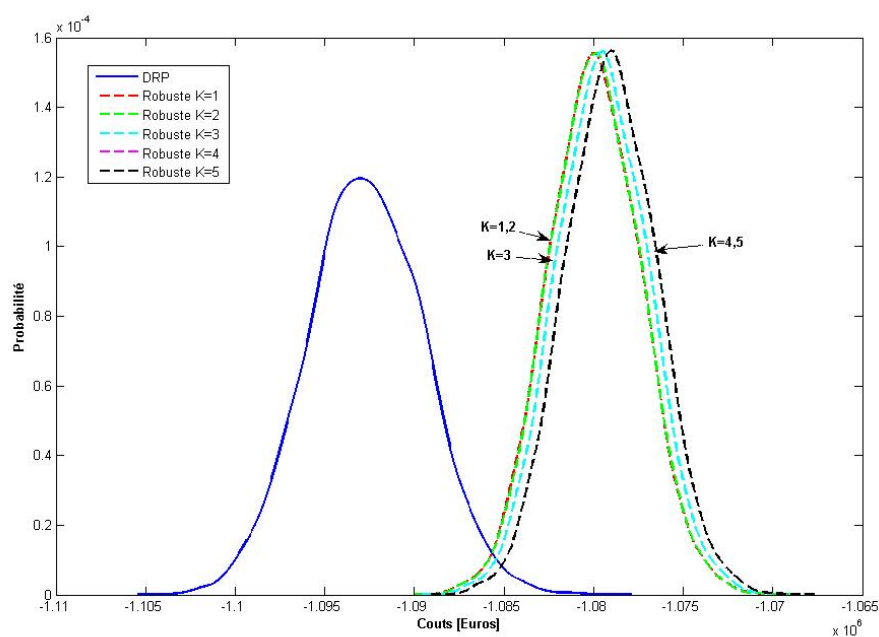
TAB. A.3 – Coûts moyens comme fonction du nombre de termes dans la règle de décision linéaire

Il semble, à la vue des résultats, que, quelle que soit la couverture, ne considérer qu'un seul terme dans la règle est suffisant. Cette conclusion permet d'éviter un accroissement inutile de la taille du modèle robuste.

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS



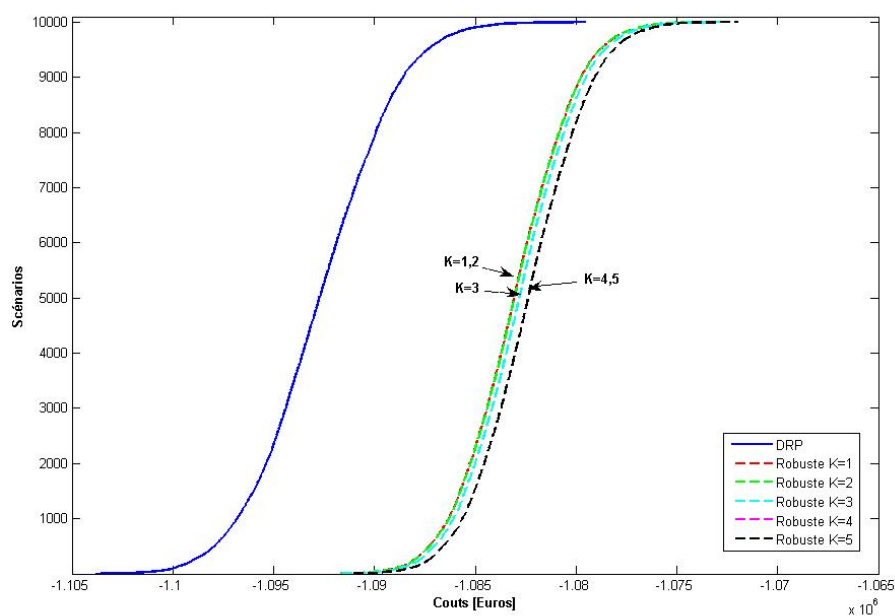
(a)



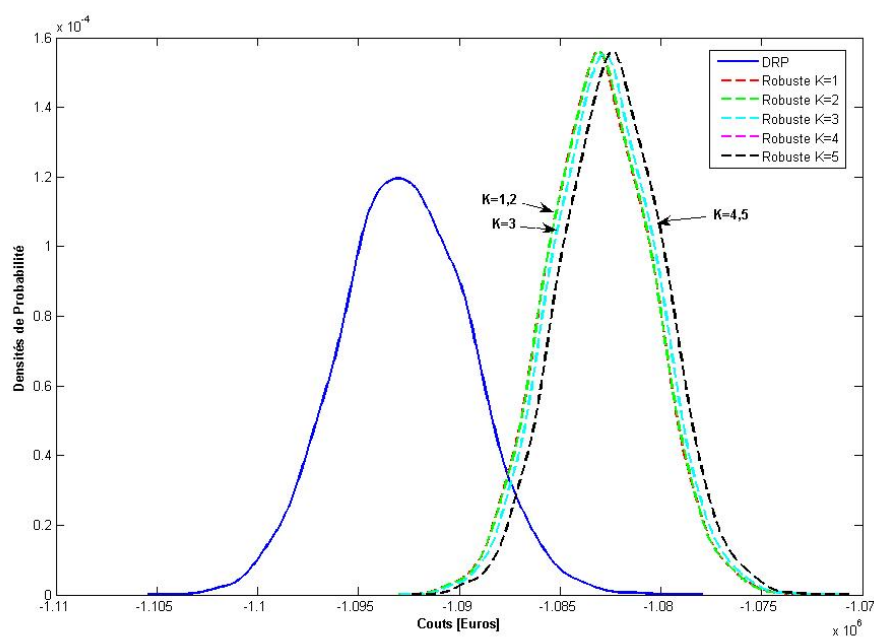
(b)

FIG. A.2 – Distribution (a) et densités (b) de probabilité des coûts du modèle robuste (couverture 90% et objectif à 0%) comme fonctions du nombre de termes K dans la règle

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS



(a)



(b)

FIG. A.3 – Distribution (a) et densités (b) de probabilité des coûts du modèle robuste (couverture 50% et objectif à 0%) comme fonctions du nombre de termes K dans la règle

A.4 Comportement du modèle RAL selon le niveau de robustesse des contraintes

Dans cette section, nous figeons le nombre de termes dans la règle à 1 et nous considérons un objectif non robuste. Les résultats suivants mettent en avant le comportement du modèle robuste : plus nous nous couvrons, moins les contraintes de volume sont violées (en V_{min} et en V_{max}) mais aussi plus les coûts sont élevés.

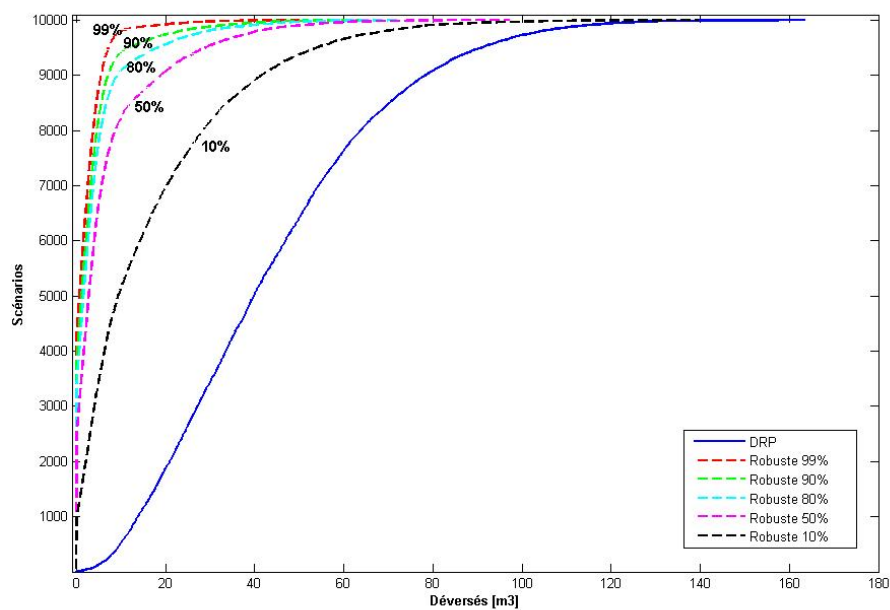
Contraintes violées	Taux de couverture					
	99%	90%	80%	50%	25%	10%
Déversés [m^3 /scen]	0,69	10,41	16,35	21,48	28,54	34,4
Δ_{Dev} (DRP) (57,2 m^3 /scen)	-98,8%	-81,8%	-71,41%	-62,45%	-50,1%	-39,9%
V_{min} [m^3 /scen]	0,395	0,395	0,395	0,395	0,404	0,463
Δ_{min} (DRP) (7,36 m^3 /scen)	-94,6%	-94,6%	-94,6%	-94,6%	-94,5%	-93,7%

TAB. A.4 – Comparatif violations de contraintes entre l'approche DRP et l'approche Robuste

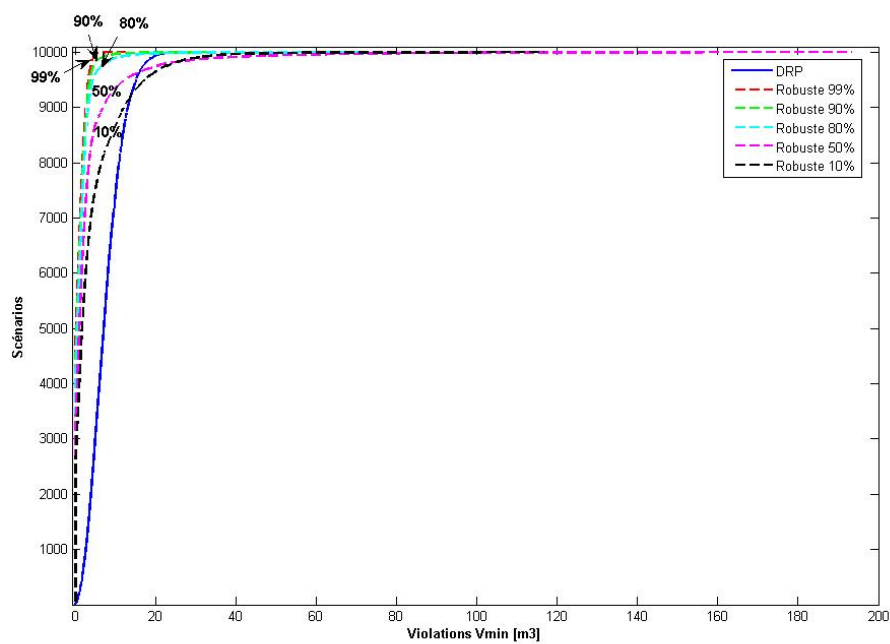
Coûts	Taux de couverture					
	99%	90%	80%	50%	25%	10 %
Coût Moyen R	-1,0763 10^6	-1,0799 10^6	-1,0811 10^6	-1,083 10^6	-1,0844 10^6	-1,0856 10^6
Δ (DRP) (-1,0926 10^6)	+1,5%	+1,16%	+1,05%	+0,88%	+0,75%	+0,64%
Ecart-Type	2,572 10^3	2,569 10^3	2,579 10^3	2,587 10^3	2,596 10^3	2,605 10^3
Coût assuré	-1,0763 10^6	-1,0799 10^6	-1,0811 10^6	-1,0830 10^6	-1,08438 10^6	-1,0855 10^6

TAB. A.5 – Comparatif des coûts en euros entre l'approche DRP à Révision Périodique (DRP) et l'approche Robuste (R)

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS



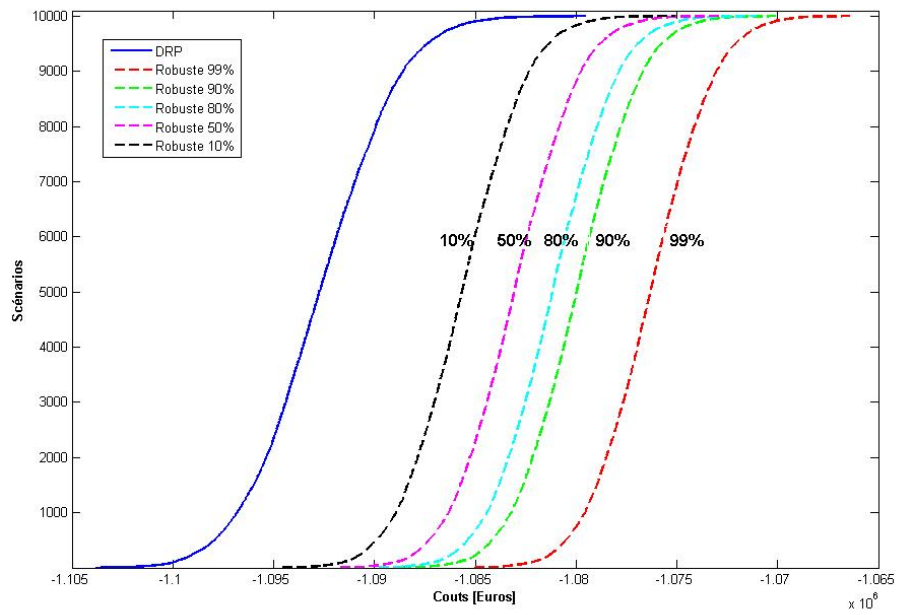
(a)



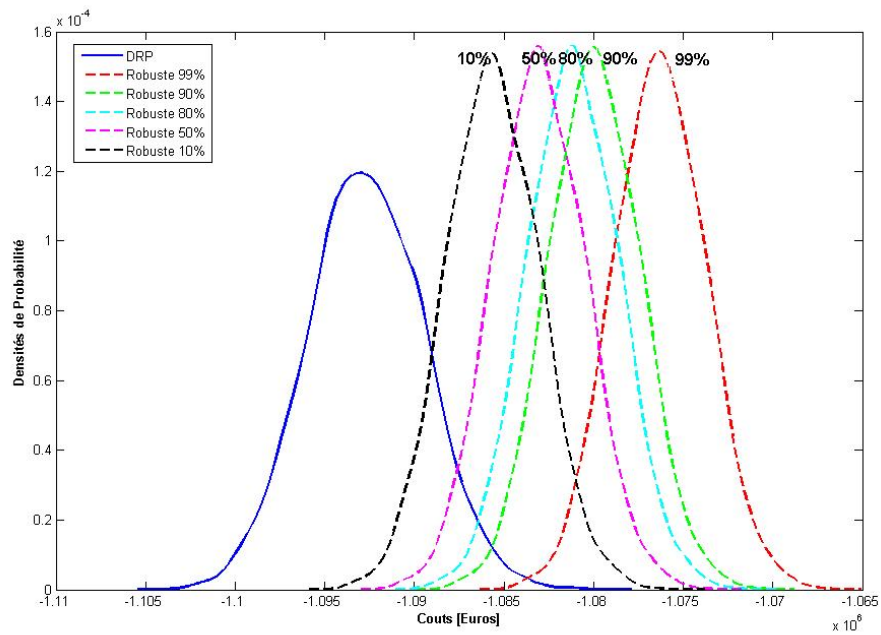
(b)

FIG. A.4 – Comparatif des fonctions de répartition des déversés cumulés par scénario (a) et des violations cumulées par scénario sur V_{min} (b) pour les modèles DRP et Robuste.

ANNEXE A. APPLICATION DE L'OPTIMISATION ROBUSTE À LA GESTION HYDRAULIQUE HEBDOMADAIRE : COMPLÉMENTS



(a)



(b)

FIG. A.5 – Comparatif des fonctions de répartition (a) et des densités de probabilité (b) des coûts moyens pour les modèles DRP et Robuste.

Annexe B

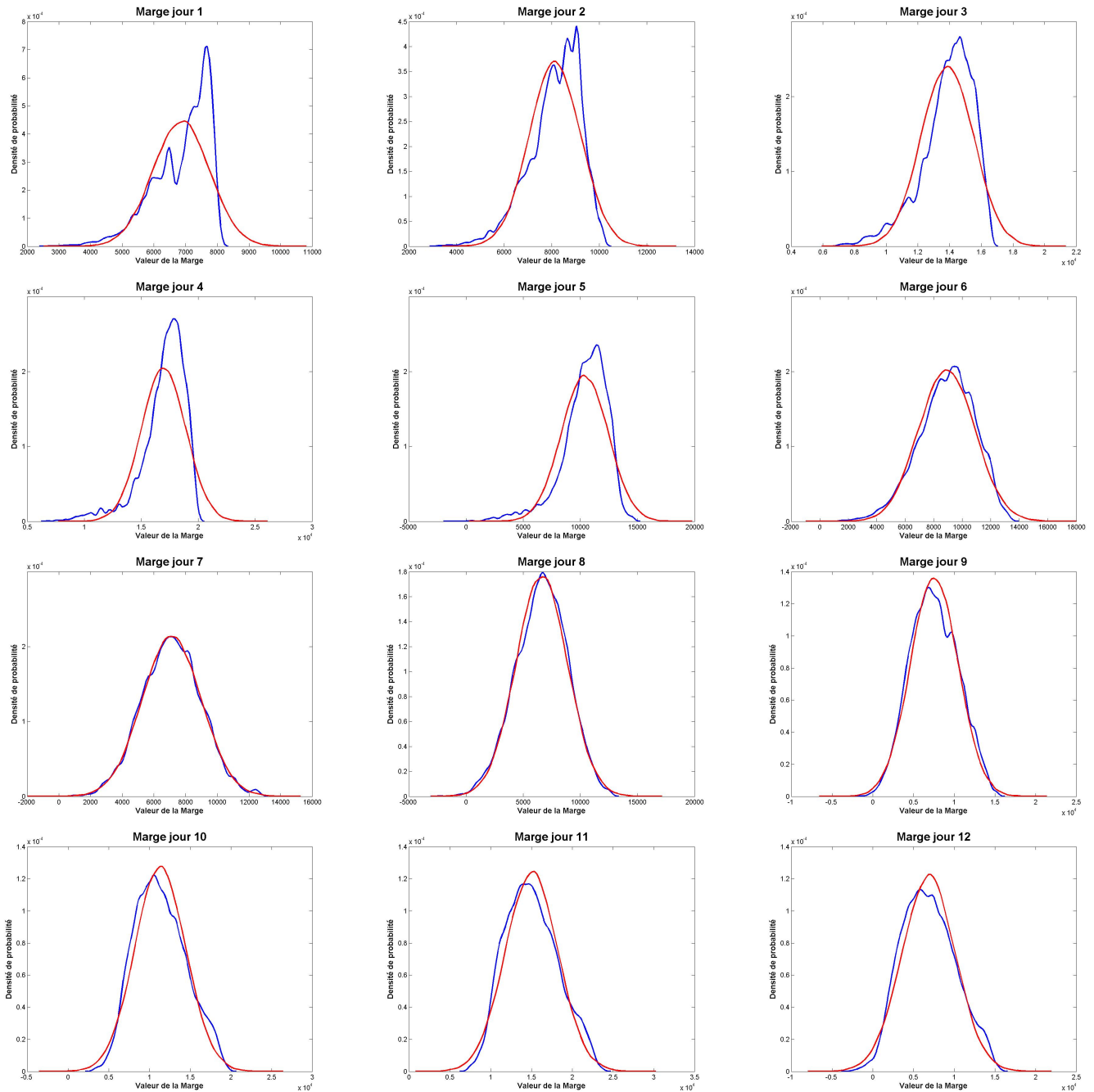
Comparatif entre les distributions de marge empiriques et une approximation par loi normale

Dans cette annexe, nous comparons :

- une distribution empirique de la marge, obtenue sur la base de 75 000 scénarios ;
- une distribution de la loi normale de moyenne et d'écart-type, ces informations étant obtenues à partir des scénarios.

Cette comparaison est utile notamment dans le cadre de l'application de l'optimisation robuste. Sur certaines journées, l'approximation par la loi normale est très satisfaisante (jours 6 à 9 par exemple). Par contre, d'autres ne le sont pas du tout, que ce soit du point de vue de la forme (jour 1) ou des queues de distribution (jours 10 à 12). Une telle approximation peut donc s'avérer insatisfaisante.

ANNEXE B. COMPARATIF ENTRE LES DISTRIBUTIONS DE MARGE EMPIRIQUES ET UNE APPROXIMATION PAR LOI NORMALE



Annexe C

Calcul de la marge physique de production : calculs complémentaires

C.1 Capacité de production maximale thermique : calcul des probabilités de panne

Nous raisonnons avec la version discrète de la loi. Si nous ne tenons pas compte des durées de pannes, la probabilité d'être en fonctionnement à l'instant n est supposée suivre une loi exponentielle de paramètre α : $\mathbb{P}(d \geq n) = e^{-\alpha n}$, $n = 1, 2, \dots$. Une panne à l'instant n signifie : $d < n$. L'apparition de la première panne à l'instant n se traduit donc par : $d \geq n - 1$ et $d < n$, ou encore : $\{d < n\} = \{d \geq n - 1\} \cap \{d < n\}$. Si F_n^1 représente l'événement "survenue de la première panne à l'instant n ", la probabilité de cet événement, notée $pp(n)$ est :

$$\begin{aligned} pp(n) &= \mathbb{P}\{F_n^1\} = \mathbb{P}\{d \geq n - 1\} - \mathbb{P}\{d < n\} \\ &= e^{-\alpha(n-1)} - e^{-\alpha n} = e^{-\alpha n}(e^\alpha - 1). \end{aligned}$$

Nous vérifions bien

$$\begin{aligned} \sum_{n=1}^{\infty} pp(n) &= (e^\alpha - 1) \sum_{n=1}^{\infty} e^{-\alpha n} = (e^\alpha - 1) \left[-1 + \frac{1}{1 - e^{-\alpha}} \right] \\ &= (e^\alpha - 1) \frac{1 - (1 - e^{-\alpha})}{1 - e^{-\alpha}} = \frac{(e^\alpha - 1)e^{-\alpha}}{1 - e^{-\alpha}} = \frac{1 - e^{-\alpha}}{1 - e^{-\alpha}} = 1. \end{aligned}$$

Au coefficient multiplicatif $(e^\alpha - 1)$ près, l'apparition de la première panne suit donc une loi exponentielle.

ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS COMPLÉMENTAIRES

Si nous tenons maintenant compte des durées de panne, le calcul de la probabilité d'être en fonctionnement à un instant donné est beaucoup plus compliqué et nécessite deux phases.

C.1.1 Phase 1 : initialisation

Nous supposons que la centrale est en fonctionnement à l'instant 0, et nous nous intéressons à l'intervalle $[1, \dots, K]$. Nous notons F_n l'évènement "la centrale est en panne à l'instant n " et f_n sa probabilité. Soit G_n l'évènement "la centrale fonctionne à l'instant n ". L'évènement $F_{n=1}$ (panne à l'instant 1), se traduit par $d < 1$, et donc $f_1 = 1 - e^{-\alpha}$. La probabilité de l'évènement F_2 , panne à l'instant $n = 2$, s'exprime en fonction de l'état antérieur de la centrale (marche ou arrêt à l'instant 1), c'est à dire en fonction de la probabilité $\mathbb{P}\{F_2|F_1\}$ d'être en panne à l'instant 2 sachant que nous l'étions à l'instant 1 ou de la probabilité $\mathbb{P}\{F_2|G_1\}$ d'être en panne à l'instant 2 sachant que la centrale fonctionnait à l'instant 1 :

$$\mathbb{P}\{F_2\} = \mathbb{P}\{F_2|F_1\}\mathbb{P}\{F_1\} + \mathbb{P}\{F_2|G_1\}\mathbb{P}\{G_1\}.$$

Explicitons les différents termes :

$\mathbb{P}\{F_2|F_1\} = 1$ car il faut un délai pour réparer ;
 $\mathbb{P}\{F_1\} = f_1$;
 $\mathbb{P}\{F_2|G_1\} = f_1$: si la centrale fonctionnait à l'instant 1, la loi exponentielle se réinitialise.
 $\mathbb{P}\{G_1\} = 1 - f_1$.

Nous avons donc : $\mathbb{P}\{F_2\} = f_1 + f_1(1 - f_1) = f_1(2 - f_1)$. D'où : $f_2 = (1 - e^{-\alpha})(1 + e^{-\alpha}) = 1 - e^{-2\alpha}$.

Pour l'instant $n = 3$, nous avons de façon similaire $\mathbb{P}\{F_3\} = \mathbb{P}\{F_3|F_2\}\mathbb{P}\{F_2\} + \mathbb{P}\{F_3|G_2\}\mathbb{P}\{G_2\}$ avec :

$\mathbb{P}\{F_3|F_2\} = 1$
 $\mathbb{P}\{F_2\} = f_2$
 $\mathbb{P}\{F_3|G_2\} = f_1$
 $\mathbb{P}\{G_2\} = 1 - f_2$

d'où : $f_3 = f_2 + f_1(1 - f_2)$.

Plus généralement, pour tout $k \leq m$, nous pouvons écrire :

$$\mathbb{P}\{F_k\} = \mathbb{P}\{F_k|F_{k-1}\}\mathbb{P}\{F_{k-1}\} + \mathbb{P}\{F_k|G_{k-1}\}\mathbb{P}\{G_{k-1}\}, \quad (\text{C.1})$$

ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS COMPLÉMENTAIRES

et nous avons :

$\mathbb{P}\{F_k|F_{k-1}\} = 1$: la centrale ne peut être réparée à l'instant k , sachant qu'elle est en panne à l'instant $k-1$;
 $\mathbb{P}\{F_k|G_{k-1}\} = f_1$, d'après la propriété de réinitialisation de la loi exponentielle.

Nous obtenons donc la formule de récurrence, pour $k \leq m$: $f_k = f_{k-1} + f_1(1 - f_{k-1})$. En ré-écrivant cette formule en $f_k = f_{k-1}(1 - f_1) + f_1$, ou encore $1 - f_k = (1 - f_{k-1})(1 - f_1)$, nous avons $1 - f_k = (1 - f_1)^k$, d'où finalement :

$$f_k = 1 - (1 - f_1)^k. \quad (\text{C.2})$$

La centrale peut tomber en panne à l'instant 1 et elle sera indisponible jusqu'à l'instant $m+1$. A l'instant $m+2$, elle peut de nouveau tomber en panne. Il faut donc considérer l'instant $m+2$. Nous écrivons comme précédemment :

$$\mathbb{P}\{F_{m+2}\} = \mathbb{P}\{F_{m+2}|F_1\}\mathbb{P}\{F_1\} + \mathbb{P}\{F_{m+2}|G_1\}\mathbb{P}\{G_1\}. \quad (\text{C.3})$$

$\mathbb{P}\{F_{m+2}|F_1\}$ se calcule simplement : la centrale a redémarré à l'instant $m+2$; la probabilité qu'elle tombe en panne à l'instant suivant vaut f_1 . Nous avons $\mathbb{P}\{F_1\} = f_1$. Le terme $\mathbb{P}\{F_{m+2}|G_1\}$ représente la probabilité que la centrale soit en fonctionnement à l'instant 1. Nous avons donc $\mathbb{P}\{F_{m+2}|G_1\} = f_{m+1}$ et $\mathbb{P}\{G_1\} = 1 - f_1$. Cette dernière formule se calcule donc par récurrence.

C.1.2 Phase 2 : généralisation

Nous nous plaçons à l'instant n et nous voulons déterminer la probabilité de l'évènement "la centrale est en fonctionnement à l'instant n ", évènement noté G_n . Nous supposons calculées les probabilités des évènements G_1, \dots, G_{n-1} (nous avons vu en phase d'initialisation comment calculer $\mathbb{P}\{G_1\}, \dots, \mathbb{P}\{G_{n-1}\}$). Nous ne pouvons pas nous contenter de conditionner les probabilités par les évènements F_{n-1} et G_{n-1} : il faut remonter à la date de la dernière panne. Nous allons établir le :

Théorème C.1 *la probabilité de l'évènement G_n est donnée par la formule de récurrence :*

$$\mathbb{P}\{G_n\} = (1 - f_1) \sum_{k \in A_{n,m}} f_1^{C((n-k-1)/(m-1))} \mathbb{P}\{G_k\}$$

où $A_{n,m}$ désigne l'ensemble des entiers k de la forme $n - j(m+1) - 1$, $j = 1, 2, \dots$ et $C(x)$ désigne le plus petit entier $\geq x$.

Preuve : nous allons considérer un système complet d'évènements, notés O_k , de la manière suivante. Nous désignons par k le dernier instant de bon fonctionnement de la centrale :

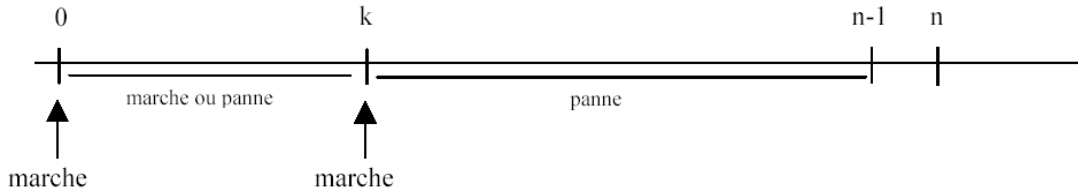
ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS COMPLÉMENTAIRES

$k = 0$: par hypothèse, la centrale fonctionnait à l'instant 0 mais a toujours été en panne ensuite (jusqu'à l'instant $n - 1$ inclus) ;

$k = 1$: la centrale fonctionnait à l'instant 0 et à l'instant 1 mais a été en panne des instants 2 à $n - 1$;

$k = 2$: la centrale fonctionnait à l'instant 2 mais a été en panne des instants $k - 1$ à $n - 1$;

$k = n - 1$: entre les instants 1 et $n - 2$, nous ne savons rien, mais à l'instant $n - 1$ la centrale fonctionnait.

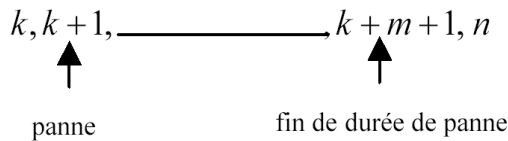


Il est évident que les O_k sont disjoints et que la réunion des O_k couvre toutes les possibilités. Pour calculer $\mathbb{P}\{G_n\}$, nous décomposons :

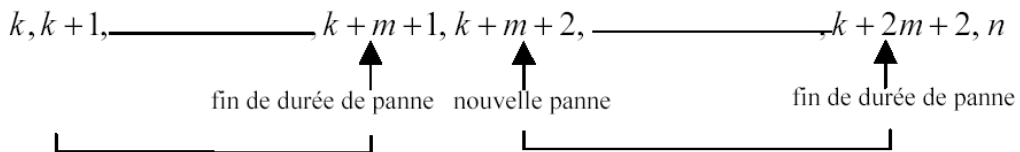
$$\mathbb{P}\{G_n\} = \sum_{k=0}^{n-1} \mathbb{P}\{G_n | O_k\} \mathbb{P}\{O_k\}. \quad (\text{C.4})$$

Si $G_n | O_k$ désigne l'évènement "sachant que la centrale marchait en k et était en panne de $k + 1$ à $n - 1$ ", la probabilité que la centrale soit de nouveau en activité à l'instant n est $1 - f_1$ si la durée de panne précédente est finie et 0 si cette durée n'est pas finie. Détaillons ces deux cas.

La durée de panne précédente est finie si l'intervalle $k + 1, \dots, n - 1$ contient un multiple entier de durées de pannes. Un intervalle d'indisponibilité est de longueur $m + 1$. Prenons par exemple le cas $n = k + m + 2$:



ou bien le cas avec $n = k + 2m + 3$:



ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS COMPLÉMENTAIRES

Plus généralement, ceci n'est possible que si $n = k + j(m + 1) + 1$, $j = 1, 2, \dots$, c'est à dire si $k \in A_{n,m}$, ensemble défini dans l'énoncé du théorème. En ce cas : $\mathbb{P}\{G_n|O_k\} = 1 - f_1$. Dans tous les autres cas : $\mathbb{P}\{G_n|O_k\} = 0$.

Nous allons calculer la probabilité que la centrale, en fonctionnement à l'instant 0, soit en panne pendant un intervalle de temps de longueur au moins N . Nous aurons besoin d'un lemme :

Lemme C.1 *la probabilité d'immobilisation pendant un intervalle de temps de longueur au moins N , notée β_N , vaut : $\beta_N = f_1^{C(N/m+1)}$ où $C(x)$ désigne le plus petit entier $\geq x$.*

Preuve : pour que la centrale soit immobilisée pendant un intervalle de temps de longueur $m+1$, il faut qu'elle tombe en panne tout de suite. La probabilité est f_1 . Pour un intervalle de longueur $2(m+1)$, il faut qu'elle tombe en panne immédiatement après la réparation. La probabilité est f_1^2 . Pour un intervalle de longueur $k(m+1)$, il faut k pannes : une au début puis après chaque réparation. La probabilité est f_1^k . Pour N donné, soit k tel que $k(m+1)$ soit immédiatement après N : $(k-1)(m+1) < N \leq k(m+1)$ et $k-1 < \frac{N}{m+1} \leq k$, d'où $k = C(N/m+1)$, plus petit entier $\geq \frac{N}{m+1}$. \square

Dans la formule (C.4), le calcul des $\mathbb{P}\{O_k\}$ se mène ainsi. Nous notons O_0 l'évènement "la centrale est tombée en panne à l'instant 1 et y est restée jusqu'à $n-1$ ", dont la probabilité est β_{n-1} , et O_1 l'évènement "la centrale marchait à l'instant 1 mais a été en panne des instants 2 à $n-1$ ". Plus généralement, pour O_k , nous écrivons :

$$\mathbb{P}(O_k) = \mathbb{P}\{O_k|F_k\}\mathbb{P}\{F_k\} + \mathbb{P}\{O_k|G_k\}\mathbb{P}\{G_k\}. \quad (\text{C.5})$$

Comme l'évènement $O_k|F_k$ est impossible, nous avons $\mathbb{P}\{O_k\} = \mathbb{P}\{O_k|G_k\}\mathbb{P}\{G_k\}$. La probabilité $\mathbb{P}\{G_k\}$ est déterminée par récurrence, et l'intervalle, exprimé en nombre de pas de temps, pendant lequel la centrale marchait en k , et a été en panne de $k+1$ à $n-1$ (évènement $O_k|G_k$) vaut $(n-1) - (k+1) - 1 = n - k - 1$ pas de temps. La probabilité de cet évènement est donc : $\mathbb{P}\{O_k|G_k\} = \beta_{n-k-1}$, d'où :

$$\mathbb{P}\{O_k\} = \beta_{n-k-1}\mathbb{P}\{G_k\}. \quad (\text{C.6})$$

Finalement, dans la formule de calcul de $\mathbb{P}\{G_n\}$, tous les termes sont connus par récurrence. La somme contient seulement les k tels que : $k = n-1-j(m+1)$, $j = 1, 2, \dots$, c'est à dire $k = n-1-(m+1)$, $k = n-1-2(m+1)$ et ainsi de suite. Pour tous les autres, la probabilité est nulle. Le théorème est ainsi démontré.

\square

C.2 Capacité de production maximale hydraulique : calcul des paramètres de la loi normale bi-tronquée

Par définition, la densité de probabilité d'une loi normale tronquée est proportionnelle à la densité de la loi normale avec l'inverse de la probabilité comme facteur de proportionnalité :

$$F_{\alpha\beta} = \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du.$$

La moyenne m de la loi bi-tronquée est :

$$m = \frac{1}{F_{\alpha\beta}} \int_{\alpha}^{\beta} u \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du.$$

En intégrant par parties et en faisant un changement de variable, nous obtenons :

$$\begin{aligned} m &= \mu \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du + \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} \frac{u-\mu}{\sigma^2} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du \\ &= \mu + \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{1}{2}\left(\frac{\alpha-\mu}{\sigma}\right)^2}^{\frac{1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2} (-e^{-\nu}) d\nu \\ &= \mu + \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \left[e^{-\frac{1}{2}\left(\frac{\alpha-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2} \right]. \end{aligned}$$

La variance se calcule par la formule $s^2 = \mathbb{E}(u^2) - m^2$, d'où :

$$\begin{aligned} s^2 &= \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} u^2 e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du - m^2 \\ &= \frac{1}{F_{\alpha\beta}} \frac{\mu}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} u e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du + \frac{1}{F_{\alpha\beta}} \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} u(u-\mu) e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du - m^2 \\ &= \frac{1}{F_{\alpha\beta}} \frac{\sigma}{\sqrt{2\pi}} \int_{\alpha}^{\beta} u \left(\frac{u-\mu}{\sigma^2} \right) e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du - m^2 + m\mu \end{aligned}$$

En remarquant que :

$$- \left(\frac{u-\mu}{\sigma^2} \right) e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du = d \left(e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} \right),$$

et en intégrant par parties, nous obtenons finalement :

ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS
COMPLÉMENTAIRES

$$\begin{aligned}s^2 &= \frac{1}{F_{\alpha\beta}} \frac{\sigma}{\sqrt{2\pi}} \left[-ue^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} \right]_{\alpha}^{\beta} - \frac{\sigma^2}{F_{\alpha\beta}} \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du + m(\mu - m) \\ &= \sigma^2 + m(\mu - m) + \frac{1}{F_{\alpha\beta}} \frac{\sigma}{\sqrt{2\pi}} \left[\alpha e^{-\frac{1}{2}\left(\frac{\alpha-\mu}{\sigma}\right)^2} - \beta e^{-\frac{1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2} \right].\end{aligned}$$

ANNEXE C. CALCUL DE LA MARGE PHYSIQUE DE PRODUCTION : CALCULS
COMPLÉMENTAIRES

Annexe D

Estimation de la densité par la méthode à noyau

Nous estimons la densité d'une variable aléatoire M en m par m_i , ce nombre d'occurrences de réalisations m_i appartenant à la i^{eme} classe associée à la valeur m . La densité est ainsi la même quelle que soit la position de m entre les extrémités de cette classe.

Une première amélioration de la technique de l'histogramme consiste à utiliser une fenêtre mobile. Nous construisons autour de m une classe de longueur h : $I_m = [m - \frac{h}{2}, m + \frac{h}{2}]$ et nous comptons le nouveau nombre d'occurrences appartenant à cette fenêtre :

$$\hat{f}(m) = Card \left\{ m_i : m - \frac{h}{2} \leq m < m + \frac{h}{2} \right\}.$$

Nous pouvons ainsi écrire :

$$\hat{f}_h(m) = \frac{1}{nh} \sum_{i=1}^n K \left(\frac{m - m_i}{h} \right)$$

où K est la fonction indicatrice de l'intervalle $[-1/2; 1/2[$, soit :

$$K(u) = \begin{cases} 0 & \text{si } u \geq \frac{1}{2} \text{ ou } u < -\frac{1}{2} \\ 1 & \text{si } -\frac{1}{2} \leq u < \frac{1}{2} \end{cases}$$

Ainsi, $K(u) = 1$ si $m_i \in I_m$.

Comme cette technique donne une estimation peu régulière, nous effectuons une seconde amélioration qui consiste à généraliser la technique précédente par introduction de fonctions K plus

continues : les noyaux. Ils permettent d'obtenir des densités plus régulières. En général, nous utilisons souvent des noyaux réguliers et symétriques, comme le noyau gaussien.

En statistique non paramétrique, nous disposons d'un certain nombre de résultats et de critères de qualité pour apprécier la qualité de l'estimation de la densité. Ainsi, pour étudier la convergence vers f d'une suite d'estimateurs $(\hat{f}_n, n \geq 1)$ de f , toutes les deux supposées être des fonctions bornées, nous définissons la distance d_{sup} de la convergence uniforme par :

$$d_{\text{sup}}(f, \hat{f}_n) = \sup_{m \in \mathbb{R}} |f(m) - \hat{f}_n(m)|.$$

Nous espérons ainsi, que si la densité f est bornée, alors la suite des distances $(d_{\text{sup}}(f, \hat{f}_n), n \geq 1)$ relatives à la suite $(\hat{f}_n, n \geq 1)$ d'estimateurs bornés de f , converge presque sûrement vers zéro.

Annexe E

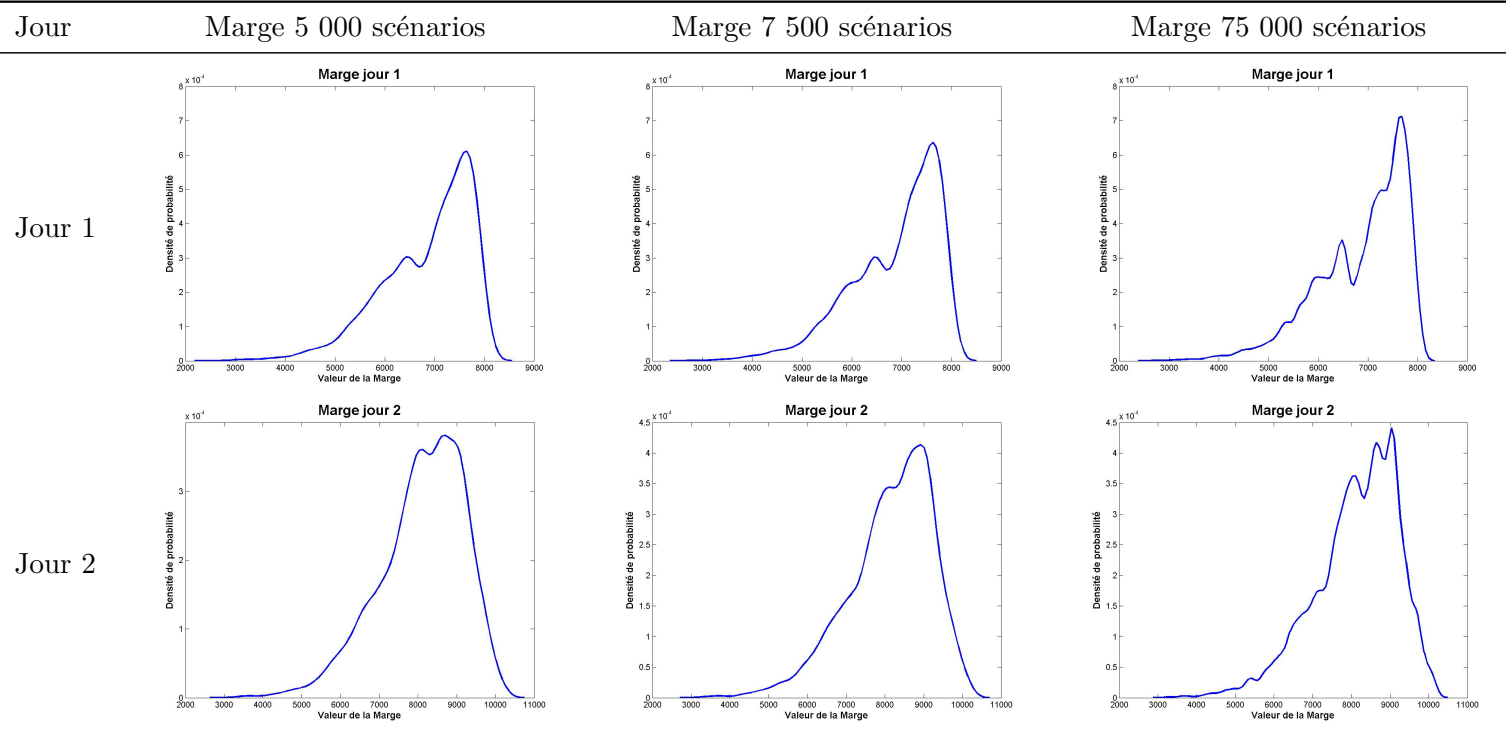
Estimation de la marge à partir de scénarios

Nous consignons dans cette annexe les résultats d'estimation de la marge prévisionnelle de production, obtenus sur données réelles (semaine du 23 mars 2006), avec la procédure décrite au chapitre 7. Les calculs ont été menés avec 5000, 7500 et 75000 scénarios.

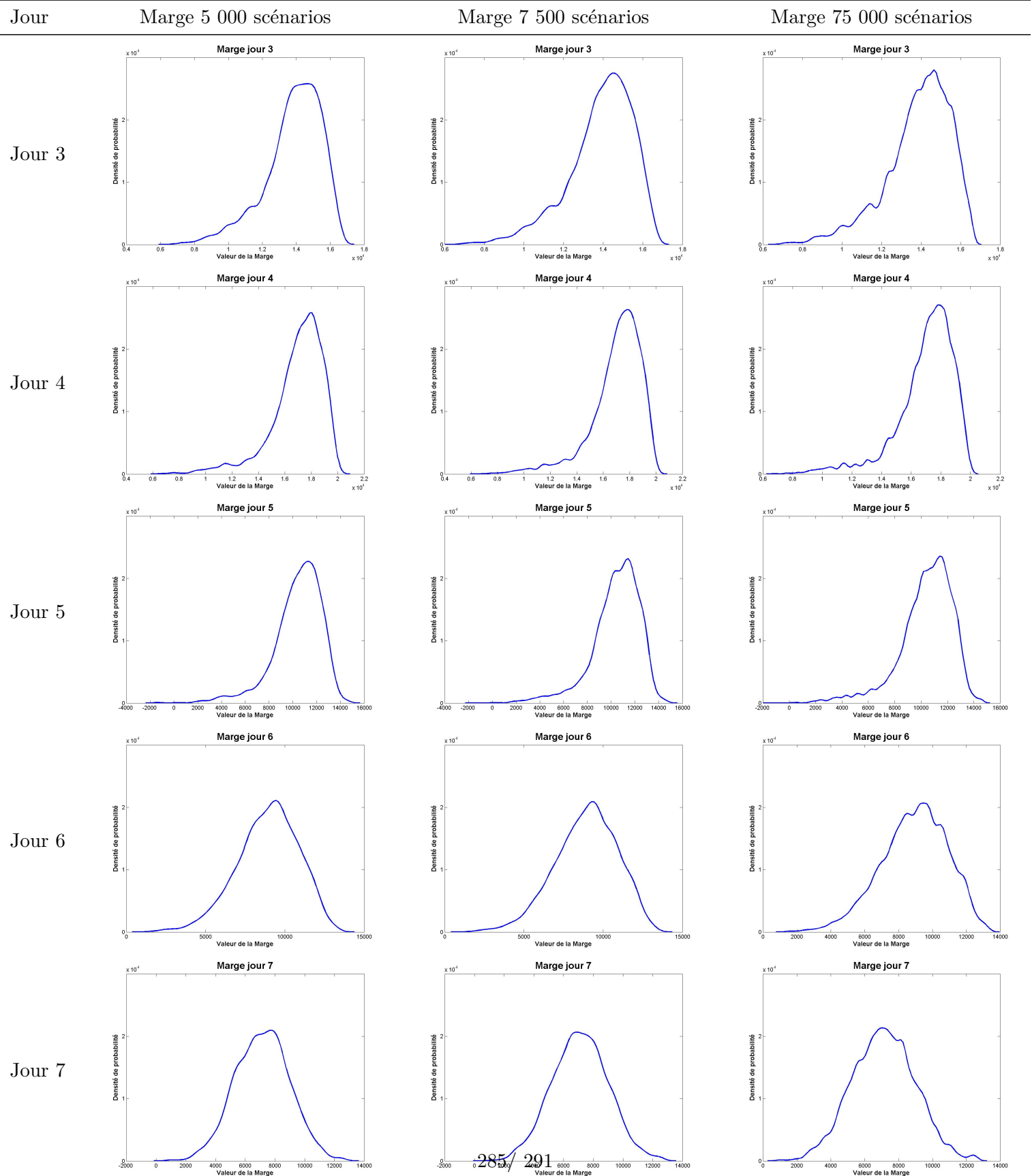
E.1 Distributions de marge

Les valeurs de marge sont exprimées en MW.

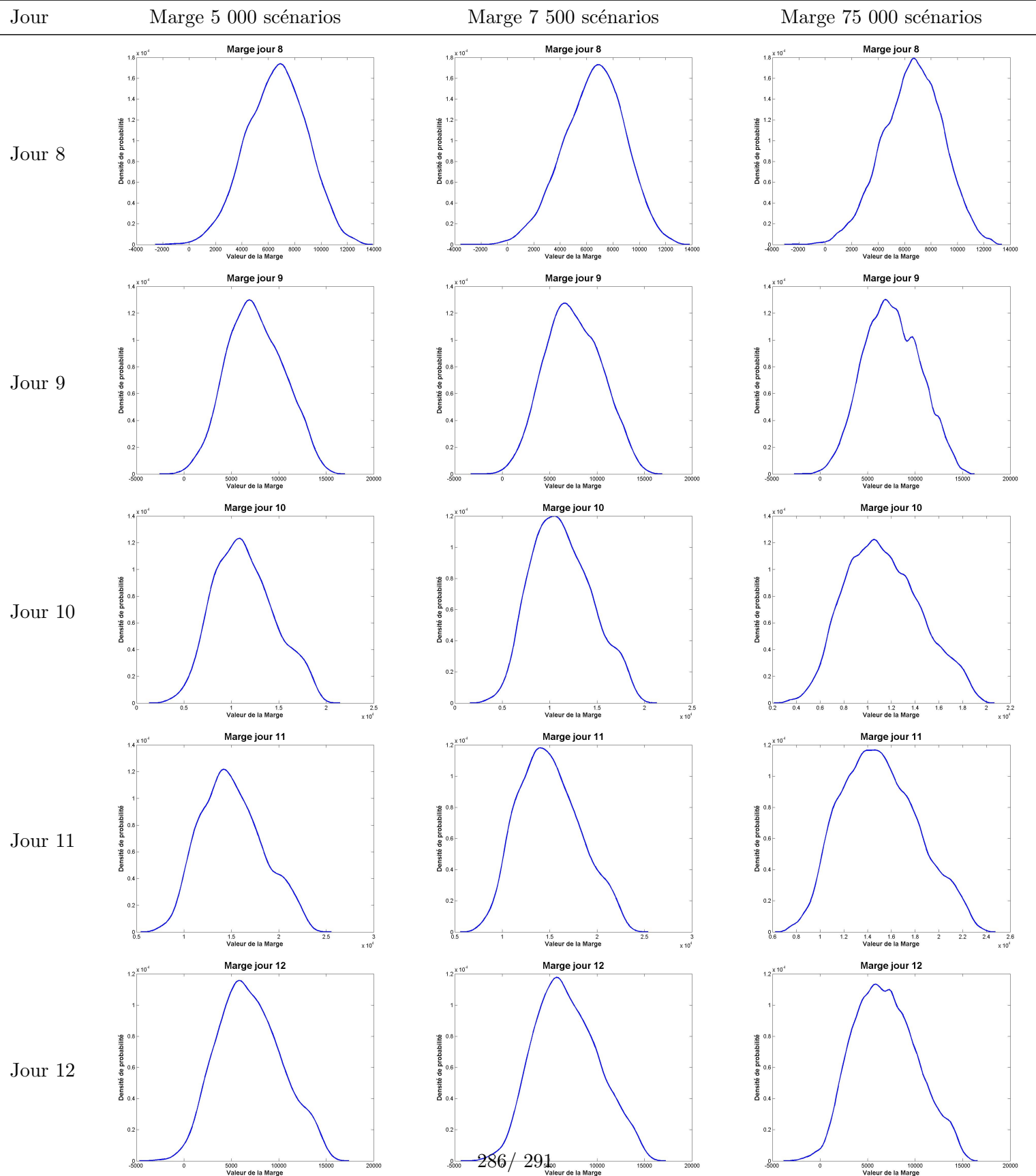
ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS



ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS



ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS



E.2 Comparaison des valeurs de quantiles pour 5 000, 7 500 et 75 000 scénarios

Nous comparons ici les valeurs de la marge extraites à différents quantiles pour les distributions obtenues avec 5 000, 7 500 et 75 000 scénarios. Au vu des résultats obtenus, il est clair que 5 000 scénarios fournissent des résultats proches de ceux obtenus avec 75 000 scénarios. Les valeurs de marge sont exprimées en MW.

JOUR 1	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	4222.29	4196.46	4208.05
Quantile 2,5%	4725.66	4739.68	4744.66
Quantile 25%	6274.62	6322.20	6314.93
Quantile 50%	7092.49	7111.90	7103.81
Quantile 75%	7591.75	7594.79	7601.25
Quantile 97,5%	7955.63	7965.03	7958.61
Min	2692.96	2806.33	2675.89
Max	8040.64	8041.15	8051.00
Moyenne	6848.00	6866.62	6865.73

JOUR 2	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	4981.20	4990.64	4980.04
Quantile 2,5%	5659.10	5684.99	5646.40
Quantile 25%	7509.42	7544.82	7544.39
Quantile 50%	8213.70	8298.00	8285.05
Quantile 75%	8938.79	8983.11	8970.73
Quantile 97,5%	9716.69	9749.12	9735.46
Min	3252.47	3277.48	3235.76
Max	10129.97	10129.97	10135.53
Moyenne	8099.74	8137.95	8131.14

ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS

JOUR 3	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	8762.72	8918.42	8802.60
Quantile 2,5%	9715.44	9908.11	9812.51
Quantile 25%	13014.48	13013.78	13014.65
Quantile 50%	14103.08	14164.22	14154.19
Quantile 75%	15055.78	15069.72	15073.83
Quantile 97,5%	16256.93	16272.25	16266.84
Min	6774.16	6806.84	6769.89
Max	16531.83	16532.39	16544.07
Moyenne	13825.09	13872.34	13859.14

JOUR 4	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	9984.63	10038.27	9946.60
Quantile 2,5%	11474.29	11494.36	11384.66
Quantile 25%	16133.29	16217.76	16175.03
Quantile 50%	17321.76	17373.91	17357.22
Quantile 75%	18292.14	18329.57	18286.55
Quantile 97,5%	19535.18	19544.56	19528.94
Min	6752.74	6753.77	6737.54
Max	20019.33	20012.83	20020.93
Moyenne	16963.98	17002.56	16979.17

JOUR 5	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	3660.34	3519.02	3479.60
Quantile 2,5%	5038.79	5003.55	4952.08
Quantile 25%	9439.97	9441.02	9427.91
Quantile 50%	10722.48	10690.66	10695.19
Quantile 75%	11816.38	11805.70	11794.06
Quantile 97,5%	13291.89	13249.05	13259.97
Min	-1352.64	-1342.04	-1377.58
Max	14615.96	14626.22	14631.29
Moyenne	10414.14	10405.68	10392.85

ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS

JOUR 6	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	3909.90	3891.95	3902.53
Quantile 2,5%	4776.91	4738.37	4732.33
Quantile 25%	7697.90	7639.54	7658.83
Quantile 50%	9090.17	9075.53	9077.06
Quantile 75%	10315.72	10355.58	10351.19
Quantile 97,5%	12182.49	12336.30	12206.57
Min	1472.17	1465.32	1443.72
Max	13269.97	13275.02	13278.64
Moyenne	8922.00	8914.83	8912.77

JOUR 7	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	2892.88	2804.22	2862.17
Quantile 2,5%	3583.05	3410.21	3514.84
Quantile 25%	5856.54	5821.67	5834.79
Quantile 50%	7176.10	7104.02	7119.37
Quantile 75%	8345.73	8345.07	8344.20
Quantile 97,5%	10707.41	10850.30	10779.06
Min	916.73	916.81	910.93
Max	12574.58	12566.68	12579.05
Moyenne	7143.37	7107.00	7120.08

JOUR 8	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	1459.98	1044.81	1108.06
Quantile 2,5%	2130.17	1808.93	1959.38
Quantile 25%	5064.45	5038.08	5077.25
Quantile 50%	6713.66	6682.97	6685.61
Quantile 75%	8183.89	8147.97	8174.26
Quantile 97,5%	10827.50	10762.15	10823.31
Min	-1255.34	-2331.96	-2331.96
Max	12582.88	12585.99	12590.11
Moyenne	6630.13	6566.18	6601.11

ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS

JOUR 9	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	1245.08	1254.96	1307.95
Quantile 2,5%	2069.29	2058.56	2178.10
Quantile 25%	5387.17	5384.98	5421.83
Quantile 50%	7352.42	7338.99	7445.64
Quantile 75%	9675.21	9613.60	9681.69
Quantile 97,5%	13290.80	13117.86	13274.05
Min	-770.87	-1653.37	-1678.87
Max	15153.93	15160.05	15168.11
Moyenne	7550.90	7479.58	7552.79

JOUR 10	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	4988.56	5136.83	5097.67
Quantile 2,5%	5844.71	5912.08	5914.40
Quantile 25%	8971.73	8929.57	8956.26
Quantile 50%	11133.26	11025.97	11109.75
Quantile 75%	13454.97	13436.60	13491.07
Quantile 97,5%	18004.18	17808.85	17833.14
Min	3250.36	3357.15	3247.69
Max	19518.65	19509.85	19528.36
Moyenne	11380.41	11292.68	11343.81

JOUR 11	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	8770.28	8730.22	8764.28
Quantile 2,5%	9638.40	9510.96	9585.44
Quantile 25%	12664.46	12579.50	12597.41
Quantile 50%	14757.08	14728.03	14813.05
Quantile 75%	17249.36	17185.61	17272.60
Quantile 97,5%	21759.15	21680.10	21686.87
Min	7402.54	7399.07	7389.16
Max	23563.50	23563.50	23570.03
Moyenne	15081.97	14983.49	15040.44

ANNEXE E. ESTIMATION DE LA MARGE À PARTIR DE SCÉNARIOS

JOUR 12	5 000 scénarios	7 500 scénarios	75 000 scénarios
Quantile 1%	118.89	53.64	267.76
Quantile 2,5%	1031.35	1053.22	1142.23
Quantile 25%	4508.38	4385.14	4423.78
Quantile 50%	6658.54	6538.88	6671.13
Quantile 75%	9242.27	9104.44	9176.78
Quantile 97,5%	13706.44	13625.93	13646.04
Min	-2682.21	-2517.65	-2671.76
Max	15364.73	15368.74	15390.70
Moyenne	6940.80	6800.10	6870.65
